

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE  
REQUEST FOR FILING APPLICATION UNDER 37 CFR 1.53(b)  
WITHOUT FILING FEE OR EXECUTED INVENTOR'S DECLARATION

Assistant Commissioner for Patents  
Washington, DC 20535 PTO

Atty. Dkt. 620-125

Date: November 28, 2000

Sir:

11/28/00

This is a request for filing a new PATENT APPLICATION under Rule 53(b) entitled:  
POLYNUCLEOTIDE AND ITS USE FOR MODULATING A DEFENCE RESPONSE IN PLANTS  
without a filing fee and/or without an executed inventor's oath/declaration.

This application is made by the below identified inventor(s). Attached hereto are the following papers:

- ☐ An abstract together with  
**120** pages of specification and claims including  
**73** numbered claims and also attached is/are  
**28** sheets of accompanying drawings.  
☒ This application is based on the following prior foreign application(s):

Application No.	Country	Filing Date
9615879.5	Great Britain	29 July 1996
9622626.1	Great Britain	30 October 1996
9704789.8	Great Britain	07 March 1997
PCT/GB97/02046	PCT	29 July 1997

respectively, the entire content of which is hereby incorporated by reference in this application, and priority is hereby claimed therefrom.

- ☐ This application is based on the following prior provisional application(s):
- | Application No. | Filing Date |
|-----------------|-------------|
|-----------------|-------------|

respectively, the entire content of which is hereby incorporated by reference in this application, and priority is hereby claimed therefrom.

- ☐ Certified copy/ies of foreign applications attached.  
☒ This application is a ☒ continuation/☐ division/☐ continuation-in-part of application Serial No. 09/230,728, filed January 29, 1999; which is a 371 of PCT/GB97/02046, filed July 29, 1997.  
☒ Please amend the specification by inserting before the first line: --This application is a ☒ continuation/☐ division/☐ continuation-in-part of application Serial No. 09/230,728, filed January 29, 1999; which is a 371 of PCT/GB97/02046, filed July 29, 1997, the entire content of which is hereby incorporated by reference in this application.--  
☒ The Examiner's attention is directed to the prior art cited in the parent application by applicant and/or Examiner for the reasons stated therein. The Examiner is requested to initial and return the attached PTO-1449 Form listing that art.

Inventor: Paul M.J. Schulze-Lefert German  
(first) MI (last) (citizenship)  
Residence: (city) Norwich (state/country) Great Britain  
Post Office Address: Sainsbury Laboratory, Norwich Research Park, Colney Lane, Norwich, Great Britain  
(Zip Code) NR4 7UH

2. Inventor: Ralph Panstruga German  
(first) MI (last) (citizenship)  
Residence: (city) Aachen (state/country) Germany  
Post Office Address: Weberstrasse 35, Aachen, Germany  
(Zip Code) 52065

NOTE: FOR ADDITIONAL INVENTORS, check box ☒ and attach sheet with same information.

Address all future communications to NIXON & VANDERHYE P.C., 1100 North Glebe Road, 8<sup>th</sup> Floor, Arlington, Virginia 22201.  
1100 North Glebe Road, 8<sup>th</sup> Floor  
Arlington, Virginia 22201-4714  
Telephone: (703) 816-4000  
Facsimile: (703) 816-4100  
MJW:tat

NIXON & VANDERHYE P.C.

By Atty: Mary J. Wilson, Reg. No. 32,955

Signature:

Mary J. Wilson

JC825 U.S. PTO  
09/722377  
11/28/00

# ***U.S. PATENT APPLICATION***

***Inventor(s):*** Paul M.J. Schulze-Lefert  
Ralph Panstruga  
Rainer Buschges

***Invention:*** POLYNUCLEOTIDE AND ITS USE FOR MODULATING A DEFENCE  
RESPONSE IN PLANTS

***NIXON & VANDERHYE P.C.  
ATTORNEYS AT LAW  
1100 NORTH GLEBE ROAD  
8<sup>TH</sup> FLOOR  
ARLINGTON, VIRGINIA 22201-4714  
(703) 816-4000  
Facsimile (703) 816-4100***

## ***SPECIFICATION***

## POLYNUCLEOTIDE AND ITS USE FOR MODULATING A DEFENCE RESPONSE IN PLANTS

The present invention relates to stimulating a defence response in plants, with a view to providing the plants with enhanced pathogen resistance. More specifically, it has resulted from cloning of the barley *Mlo* gene, various mutant *mlo* alleles, and a number of homologues from various species. The *Mlo* gene has been isolated using a positional cloning approach which has never previously been successful in Barley. Details and discussion are provided below. Wild-type *Mlo* exerts a negative regulatory function on a pathogen defence response, such that mutants exhibit a defence response in the absence of pathogen. In accordance with the present invention, down-regulation or out-competition of *Mlo* function may be used to stimulate a defence response in transgenic plants, conferring increased pathogen resistance.

Mutations have been described in several plants in which defence responses to pathogens appear to be constitutively expressed. Mutation-induced recessive alleles (*mlo*) of the barley *Mlo* locus exhibit a leaf lesion phenotype and confer an apparently durable, broad spectrum resistance to the powdery mildew pathogen, *Erysiphe graminis* f sp *hordei*.

Resistance responses to the powdery mildew pathogen have been genetically well characterized (Wiberg, 1974; Søgaard and Jørgensen, 1988; Jørgensen, 1994). In most analyzed cases resistance is specified by race-specific resistance genes following the rules of Flor's gene-for-gene hypothesis (Flor, 1971). In this type of plant/pathogen interaction, resistance

is specified by and dependent on the presence of two complementary genes, one from the host and one from the fungal pathogen. The complementary genes have been termed operationally (pathogen) resistance ("R") gene and avirulence gene, respectively. Most of the powdery mildew resistance genes (Mlx) act as dominant or semidominant traits (Jørgensen, 1994).

Monogenic resistance mediated by recessive (mlo) alleles of the Mlo locus is different. Apart from being recessive, it differs from race-specific resistance to single pathogen strains in that (i) it confers broad spectrum resistance to almost all known isolates of the pathogen (ii) mlo resistance alleles have been obtained by mutagen treatment of any tested susceptible wild type (Mlo) variety, and (iii) mlo resistance alleles exhibit a defence mimic phenotype in the absence of the pathogen (Wolter et al. , 1993). Thus, the genetic data indicate the Mlo wild type allele exerts a negative regulatory function on defence responses to pathogen attack.

Resistance mediated by mlo alleles is currently widely used in barley breeding and an estimated 10 million hectares are annually planted in Europe with seeds of this genotype. A 'mlo like' inherited resistance to powdery mildew in other cereal plants has not been reported so far although the fungus is a relevant pathogen in wheat (attacked by *Erysiphe graminis* f sp *tritici*), oat (attacked by *E. g.* f sp *avenae*), and rye (attacked by *E. g.* f sp *secalis*). Because cereals are morphologically, genetically and biochemically highly related to each other (Moore et al., 1995), one would predict the

existence of homologous genes in these species. The failure to have found a 'mlo like' inherited resistance in wheat and oat is probably due to their hexaploid genomes, making it difficult to obtain by mutagenesis defective alleles in all six gene copies, and the chance of all such mutations occurring in Nature is remote. The failure to have found a *mlo* equivalent in other cereals is probably due to insignificant amount of mutational analysis in these species and complications as a result of their outbreeding nature (e.g. rye).

RFLP markers closely linked to *Mlo* on barley chromosome 4 were previously identified on the basis of a *mlo* backcross line collection containing *mlo* alleles from six genetic backgrounds (Hinze et al. , 1991). The map position of *Mlo* on the basis of RFLP markers was consistent with its chromosomal localization as determined by a previous mapping with morphological markers (Jørgensen, 1977).

Having identified an ~3cM genetic interval containing *Mlo* bordered by genetic markers, we decided to attempt to isolate the gene via positional cloning.

However, there is no documented example of a successful positional cloning attempt of a barley gene. We were faced with a number of difficulties.

Firstly, the genome of barley ( $5.3 \times 10^9$  bp/haploid genome equivalent; Bennett and Smith, 1991) has almost double the size of the human genome and because the total genetic map covers ~1.800 cM (Becker et al., 1995) we were confronted with a very unfavourable ratio of genetic and physical distances (1 cM

corresponds to ~ 3 Mb).

Secondly, a high resolution genetic map had to be constructed around *Mlo* enabling the positioning of linked markers with a precision of better than 0.1 cM.

5        Thirdly, we aimed to physically delimit the target gene and both flanking DNA markers on individual large insert genomic clones, a procedure later termed "chromosome landing" (Tanksley et al., 1995). For this purpose, a complete barley YAC library from barley Megabase DNA had to be constructed with  
10       an average insert size of 500-600 kb, which was unprecedented.

         Fourthly, we had to prepare unusual generic tools that enabled us to identify the *Mlo* gene within a physically delimited region without the need for a time consuming  
15       generation of barley transgenic plants and testing of different candidate genes. We used for our studies ten characterized radiation- or chemically-induced *mlo* mutants (Jørgensen, 1992). For a conclusive chain of evidence of the gene isolation we decided to depend upon a functional restoration of the wild  
20       type *Mlo* allele starting out from characterized *mlo* defective alleles. For this purpose, we performed *mlo* heteroallelic crosses and isolated susceptible intragenic *Mlo* recombinants. The sequence analysis of these proves the function of the described gene.

         The cloning of the barley *Mlo* gene and homologues,  
25       including homologues from other plant species, gives rise to a number of practical applications, reflected in the various aspects of the present invention.

09722377 "112800

According to a first aspect of the present invention there is provided a nucleic acid molecule comprising a nucleotide sequence encoding a peptide with *Mlo* function. Those skilled in the art will appreciate that "*Mlo* function" refers to the ability to suppress a defence response, said defence response being race and/or pathogen independent and autonomous of the presence of a pathogen, such as, for example, the *Mlo* gene of barley, the *Acd* gene and the *Lsd* gene of *Arabidopsis*.

*mlo* mutations that down-regulate or disrupt functional expression of the wild-type *Mlo* sequence are recessive, such that they are complemented by expression of a wild-type sequence. Thus "*Mlo* function" can be determined by assessing the level of constitutive defence response and/or susceptibility of the plant to a pathogen such as, for example, powdery mildew or rust (e.g. yellow rust). Accordingly, a putative nucleotide sequence with *Mlo* function can be tested upon complementation of a suitable *mlo* mutant. The term "*mlo* function" is used to refer to sequences which confer a *mlo* mutant phenotype on a plant.

The capitalisation of "*Mlo*" and non-capitalisation of "*mlo*" is thus used to differentiate between "wild-type" and "mutant" function.

A *mlo* mutant phenotype is characterised by the exhibition of an increased resistance against one or more pathogens, which is race and/or pathogen independent and autonomous of the presence of a pathogen.

The test plant may be monocotyledonous or dicotyledonous.

Suitable monocots include any of barley, rice, wheat, maize or oat, particularly barley. Suitable dicots include *Arabidopsis*.

Nucleic acid according to the invention may encode a polypeptide comprising the amino acid sequence shown in Figure 2, or an allele, variant, derivative or mutant, or homologue, thereof.

Nucleic acid according to the present invention may have the sequence of a *Mlo* gene of barley, or be a mutant, variant (or derivative) or allele of the sequence provided, or a homologue thereof. Preferred mutants, variants and alleles are those which encode a sequence which retains a functional characteristic of the wild-type gene, especially the ability to suppress a defence response as discussed herein. Other preferred mutants, variants and alleles encode a sequence which, in a homozygote, cause constitutive activation of a defence response, or at least promotes activation of a defence response (i.e. is a *mlo* mutant sequence), e.g. by reducing or wholly or partly abolishing *Mlo* function. Preferred mutations giving *mlo* mutant sequences are shown in Table 1. Changes to a sequence, to produce a mutant, derivative or variant, may be by one or more of addition, insertion, deletion or substitution of one or more nucleotides in the nucleic acid, leading to the addition, insertion, deletion and/or substitution of one or more amino acids. Of course, changes to the nucleic acid which make no difference to the encoded amino acid sequence are included. Particular variants, mutants, alleles and



derivatives are discussed further below, as well as homologues.

A preferred nucleic acid sequence according to an aspect of the present invention is shown in Figure 2 along with the predicted amino acid sequence. Nucleic acid may be subject to alteration by way of substitution of nucleotides and/or a combination of addition, insertion and/or substitution of one or more nucleotides with or without altering the encoded amino acids sequence (by virtue of the degeneracy of the genetic code).

As discussed below, further aspects of the present invention provide homologues of the *Mlo* sequence shown in Figure 2, including from rice (genomic sequence Figure 5, bottom line, cDNA sequence Figure 10, amino acid sequence Figure 13) and barley (genomic sequence Figure 6, bottom line, cDNA sequence Figure 11, amino acid sequence Figure 14); also Table 5B (nucleotide sequences) and Figure 5A (amino acid sequences) show homologous EST's from rice and *Arabidopsis*.

The present invention also provides a vector which comprises nucleic acid with any one of the provided sequences, preferably a vector from which a product can be expressed. The vector is preferably suitable for transformation into a plant cell and/or a microbial cell. The invention further encompasses a host cell transformed with such a vector, especially a plant cell or a microbial cell (e.g. *Agrobacterium tumefaciens*).

Thus, a host cell, such as a plant cell, comprising nucleic acid according to the present invention is provided. Within the cell, the nucleic acid may be incorporated within the nuclear

genome, i.e. a chromosome. There may be more than one heterologous nucleotide sequence per haploid genome.

A vector comprising nucleic acid according to the present invention need not include a promoter, particularly if the  
5 vector is to be used to introduce the nucleic acid into cells for recombination into the genome.

Nucleic acid molecules and vectors according to the present invention may be provided in a form isolated and/or purified from their natural environment, in substantially pure  
10 or homogeneous form, or free or substantially free of nucleic acid or genes of the species of interest or origin other than the relevant sequence. Nucleic acid according to the present invention may comprise cDNA, RNA, genomic DNA and may be wholly or partially synthetic. The term "isolate" may encompass all  
15 these possibilities.

The present invention also encompasses the expression product of any of the nucleic acid sequences disclosed and methods of making the expression product by expression from encoding nucleic acid therefore under suitable conditions in  
20 suitable host cells, e.g. *E. coli*. Those skilled in the art are well able to construct vectors and design protocols for expression and recovery of products of recombinant gene expression. Suitable vectors can be chosen or constructed, containing one or more appropriate regulatory sequences,  
25 including promoter sequences, terminator fragments, polyadenylation sequences, enhancer sequences, marker genes and other sequences as appropriate. For further details see, for

00722377 44300

example, *Molecular Cloning: a Laboratory Manual*: 2nd edition, Sambrook et al, 1989, Cold Spring Harbor Laboratory Press. Transformation procedures depend on the host used, but are well known. Many known techniques and protocols for manipulation of nucleic acid, for example in preparation of nucleic acid constructs, mutagenesis, sequencing, introduction of DNA into cells and gene expression, and analysis of proteins, are described in detail in *Short Protocols in Molecular Biology*, Second Edition, Ausubel et al. eds., John Wiley & Sons, 1992.

10 The disclosures of Sambrook et al. and Ausubel et al. are incorporated herein by reference, along with all other documents mentioned.

Purified Mlo protein, or a fragment, mutant or variant thereof, e.g. produced recombinantly by expression from encoding nucleic acid therefor, may be used to raise antibodies employing techniques which are standard in the art. Antibodies and polypeptides comprising antigen-binding fragments of antibodies may be used in identifying homologues from other species as discussed further below.

20 Methods of producing antibodies include immunising a mammal (eg human, mouse, rat, rabbit, horse, goat, sheep or monkey) with the protein or a fragment thereof. Antibodies may be obtained from immunised animals using any of a variety of techniques known in the art, and might be screened, preferably using binding of antibody to antigen of interest. For instance, Western blotting techniques or immunoprecipitation may be used (Armitage et al, 1992, *Nature* 357: 80-82).

As an alternative or supplement to immunising a mammal, antibodies with appropriate binding specificity may be obtained from a recombinantly produced library of expressed

Antibodies raised to a polypeptide or peptide can be used in the identification and/or isolation of homologous

polypeptides, and then the encoding genes. Thus, the present invention provides a method of identifying or isolating a polypeptide with Mlo or mlo function (in accordance with embodiments disclosed herein), comprising screening candidate peptides or polypeptides with a polypeptide comprising the antigen-binding domain of an antibody (for example whole antibody or a fragment thereof) which is able to bind an Mlo or mlo peptide, polypeptide or fragment, variant or variant thereof or preferably has binding specificity for such a peptide or polypeptide, such as having an amino acid sequence identified herein. Specific binding members such as antibodies and polypeptides comprising antigen binding domains of antibodies that bind and are preferably specific for a Mlo or mlo peptide or polypeptide or mutant, variant or derivative thereof represent further aspects of the present invention, as do their use and methods which employ them.

Candidate peptides or polypeptides for screening may for

instance be the products of an expression library created using nucleic acid derived from an plant of interest, or may be the product of a purification process from a natural source.

A peptide or polypeptide found to bind the antibody may be isolated and then may be subject to amino acid sequencing. Any suitable technique may be used to sequence the peptide or polypeptide either wholly or partially (for instance a fragment of a polypeptide may be sequenced). Amino acid sequence information may be used in obtaining nucleic acid encoding the peptide or polypeptide, for instance by designing one or more oligonucleotides (e.g. a degenerate pool of oligonucleotides) for use as probes or primers in hybridisation to candidate nucleic acid, or by searching computer sequence databases, as discussed further below.

A further aspect of the present invention provides a method of identifying and cloning *Mlo* homologues from plants, including species other than Barley, which method employs a nucleotide sequence derived from that shown in Figure 2. Further similar aspects employ a nucleotide sequence derived from any of the other Figures provided herein. Nucleic acid libraries may be screened using techniques well known to those skilled in the art and homologous sequences thereby identified then tested. The provision of sequence information for the *Mlo* gene of Barley and various homologues enables the obtention of homologous sequences from Barley and other plant species, as exemplified further herein.

Also, one can easily derive PCR primers based on putative

within the sequence shown in Figure 2, a single amino acid change with respect to the sequence shown in Figure 2, or 2, 3, 4, 5, 6, 7, 8, or 9 changes, about 10, 15, 20, 30, 40 or 50 changes, or greater than about 50, 60, 70, 80 or 90 changes.

- 5 In addition to one or more changes within the amino acid sequence shown in Figure 2, a mutant, allele, variant or derivative amino acid sequence may include additional amino acids at the C-terminus and/or N-terminus.

As is well-understood, homology at the amino acid level is  
10 generally in terms of amino acid similarity or identity. Similarity allows for "conservative variation", i.e. substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as arginine for lysine, glutamic for aspartic acid, or glutamine for asparagine.  
15 Similarity may be as defined and determined by the TBLASTN program, of Altschul et al. (1990) *J. Mol. Biol.* 215: 403-10, which is in standard use in the art, or, and this may be preferred, the standard program BestFit, which is part of the  
20 Wisconsin Package, Version 8, September 1994, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA, Wisconsin 53711). BestFit makes an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number  
25 of matches using the local homology algorithm of Smith and Waterman

Homology may be over the full-length of the relevant

exon sequences, which might be identified by comparison with the *Mlo* sequence provided in Figure 2 wherein exons are highlighted, and perform RT-PCR with total RNA from the plant of interest, e.g. barley and rice for the homologues shown in  
5 Figures 5 and 6, with cDNA and amino acid sequences shown in other figures herein.

The homologues whose nucleotide sequences are given and whose amino acid sequences are given or are deducible represent and provide further aspects of the present invention in  
10 accordance with those disclosed for the Barley gene shown in Figure 2.

The present invention also extends to nucleic acid encoding a *Mlo* homologue obtained using a nucleotide sequence derived from that shown in Figure 2, or the amino acid sequence  
15 shown in Figure 2. Preferably, the nucleotide sequence and/or amino acid sequence shares homology with the sequence encoded by the nucleotide sequence of Figure 2, preferably at least about 50%, or at least about 55%, or at least about 60%, or at least about 65%, or at least about 70%, or at least about 75%,  
20 or at least about 80% homology, or at least about 85% homology, or at least about 90% homology, most preferably at least about 95% homology. "Homology" in relation to an amino acid sequence may be used to refer to identity or similarity, preferably identity. High levels of amino acid identity may be limited to  
25 functionally significant domains or regions.

A mutant, allele, variant or derivative amino acid sequence in accordance with the present invention may include

sequence shown herein, or may more preferably be over a contiguous sequence of about or greater than about 20, 25, 30, 33, 40, 50, 67, 133, 167, 200, 233, 267, 300, 333, 400, 450, 500, 550, 600 or more amino acids or codons, compared with the relevant amino acid sequence or nucleotide sequence as the case may be.

The EST sequences provided herein, have on average 70% similarity and 50% identity with the Mlo amino acid sequence of Figure 2. We show that the rice homologue (Figure 5) and barley homologue (Figure 6) have an amino acid identity of 81% (amino acid sequences shown in Figure 13 and Figure 14).

In certain embodiments, an allele, variant, derivative, mutant or homologue of the specific sequence may show little overall homology, say about 20%, or about 25%, or about 30%, or about 35%, or about 40% or about 45%, with the specific sequence. However, in functionally significant domains or regions the amino acid homology may be much higher. Putative functionally significant domains or regions can be identified using processes of bioinformatics, including comparison of the sequences of homologues. Functionally significant domains or regions of different polypeptides may be combined for expression from encoding nucleic acid as a fusion protein. For example, particularly advantageous or desirable properties of different homologues may be combined in a hybrid protein, such that the resultant expression product, with Mlo or mlo function, may comprise fragments of various parent proteins.

The nucleotide sequence information provided herein, or



any part thereof, may be used in a data-base search to find homologous sequences, expression products of which can be tested for *Mlo* or *mlo* function. These may have ability to complement a *mlo* mutant phenotype in a plant or may, upon  
5 expression in a plant, confer a *mlo* phenotype.

In public sequence databases we recently identified several homologues for the sequence of Figure 2. We have already found homologues in rice and barley, and the dicot. *Arabidopsis*.

10 By sequencing homologues, studying their expression patterns and examining the effect of altering their expression, genes carrying out a similar function to *Mlo* in Barley are obtainable. Of course, mutants, variants and alleles of these sequences are included within the scope of the present  
15 invention in the same terms as discussed above for the Barley gene.

Homology between the homologues as disclosed herein, may be exploited in the identification of further homologues, for example using oligonucleotides (e.g. a degenerate pool)  
20 designed on the basis of sequence conservation.

According to a further aspect, the present invention provides a method of identifying or a method of cloning a *Mlo* homologue, e.g. from a species other than Barley, the method employing a nucleotide sequence derived from that shown in  
25 Figure 2 or that shown in any of the other Figures herein. For instance, such a method may employ an oligonucleotide or oligonucleotides which comprises or comprise a sequence or

sequences that are conserved between the sequences of Figures 2 and/or 5 and/or 6 and/or 10 and/or 11 and/or 12, or encoding an amino acid sequence conserved between Figure 2 and/or 7 and/or 13 and/or 14 and/or 15 to search for homologues. Thus, a method of obtaining nucleic acid is provided, comprising hybridisation of an oligonucleotide or a nucleic acid molecule comprising such an oligonucleotide to target/candidate nucleic acid. Target or candidate nucleic acid may, for example, comprise a genomic or cDNA library obtainable from an organism known to contain or suspected of containing such nucleic acid, either monocotyledonous or dicotyledonous. Successful hybridisation may be identified and target/candidate nucleic acid isolated for further investigation and/or use.

Hybridisation may involve probing nucleic acid and identifying positive hybridisation under suitably stringent conditions (in accordance with known techniques) and/or use of oligonucleotides as primers in a method of nucleic acid amplification, such as PCR. For probing, preferred conditions are those which are stringent enough for there to be a simple pattern with a small number of hybridisations identified as positive which can be investigated further. It is well known in the art to increase stringency of hybridisation gradually until only a few positive clones remain.

As an alternative to probing, though still employing nucleic acid hybridisation, oligonucleotides designed to amplify DNA sequences may be used in PCR reactions or other methods involving amplification of nucleic acid, using routine

097237 112500

procedures. See for instance "PCR protocols; A Guide to Methods and Applications", Eds. Innis et al, 1990, Academic Press, New York.

Preferred amino acid sequences suitable for use in the design of probes or PCR primers for some purposes are sequences conserved (completely, substantially or partly) between at least two Mlo peptides or polypeptides encoded by genes able to suppress a defence response in a plant, e.g. with any of the amino acid sequences of any of the various figures herein and/or encoded by the nucleotide sequences of any of the various figures herein.

On the basis of amino acid sequence information oligonucleotide probes or primers may be designed, taking into account the degeneracy of the genetic code, and, where appropriate, codon usage of the organism from the candidate nucleic acid is derived.

Preferably an oligonucleotide in accordance with certain embodiments of the invention, e.g. for use in nucleic acid amplification, is up to about 50 nucleotides, or about 40 nucleotides or about 30 or fewer nucleotides in length (e.g. 18, 21 or 24).

Assessment of whether or not such a PCR product corresponds to Mlo homologue genes may be conducted in various ways. A PCR band from such a reaction might contain a complex mix of products. Individual products may be cloned and each one individually screened. It may be analysed by transformation to assess function on introduction into a plant



25L            5'-GTG CAT CTG CGT GTG CGT A-3'  
 25LN           5'-GTG TGC GTA CCT GGT AGA G-3'  
 25R            5'-AAC GAC GTC TGG TGC GTG-3'  
 33             5'-TGC AGC TAT ATG ACC TTC CCC CTC-3'  
 5             37            5'-GGA CAT GCT GAT GGC TCA GA-3'  
 38            5'-CAG AAC TTG TCT CAT CCC TG-3  
 38A           5'-GGC TAT ACA TTG GGA CTA ACA-3'  
 38B           5'-CGA ATC ATC ACA TCC TAT GTT-3'  
 39            5'-GCA AGT TCG ACT TCC AC-3'  
 10           39A           5'-TCG ACT TCC ACA AGT ACA TCA-3'  
 53            5'-AGC GTA CCT GCG TAC GTA G-3'

Various primer combinations have been tested:

38/39A; 38/39; 38/33; 38/37; 38A/39A; 38B/39A; 38/25L; 38/25LN;  
 25R/25L; 25R/25LN; 25R/53.

15

Various aspects of the present invention include the  
 obtainable nucleic acid, methods of screening material, e.g.  
 cell lysate, nucleic acid preparations, for the presence of  
 nucleic acid of interest, methods of obtaining the nucleic  
 20 acid, and the primers and primer combinations given above.

The sequence information provided herein also allows the  
 design of diagnostic tests for determination of the presence of  
 a specific mlo resistance allele, or a susceptibility allele  
 25 (e.g. wild-type), in any given plant, cultivar, variety,  
 population, landrace, part of a family or other selection in a  
 breeding programme or other such genotype. A diagnostic test

may be based on determination of the presence or absence of a particular allele by means of nucleic acid or polypeptide determination.

At the nucleic acid level, this may involve hybridisation of a suitable oligo- or poly-nucleotide, such as a fragment of the Mlo gene or a homologue thereof, including any homologue disclosed herein, or any particular allele, such as an allele which gives an mlo phenotype, such as any such allele disclosed herein. The hybridisation may involve PCR designed to amplify a product from a given allelic version of mlo, with subsequent detection of an amplified product by any of a number of possible methods including but not limited to gel electrophoresis, capillary electrophoresis, direct hybridisation of nucleotide sequence probes and so on. A diagnostic test may be based on PCR designed to amplify various alleles or any allele from the Mlo locus, with a test to distinguish the different possible alleles by any of a number of possible methods, including DNA fragment size, restriction site variation (e.g. CAPS - cleaved amplified polymorphic sites) and so on. A diagnostic test may also be based on a great number of possible variants of nucleic acid analysis that will be apparent to those skilled in the art, such as use of a synthetic mlo-derived sequence as a hybridisation probe.

Broadly, the methods divide into those screening for the presence of nucleic acid sequences and those that rely on detecting the presence or absence of a polypeptide. The methods may make use of biological samples from one or more

plants or cells that are suspected to contain the nucleic acid sequences or polypeptide.

Exemplary approaches for detecting nucleic acid or polypeptides include analysing a sample from the plant or plant cell by:

(a) comparing the sequence of nucleic acid in the sample with all or part of the nucleotide sequence shown in Figure 7 to determine whether the sample from the patient contains a mutation;

(b) determining the presence in the sample of a polypeptide including the amino acid sequence shown in Figure 2 or a fragment thereof and, if present, determining whether the polypeptide is full length, and/or is mutated, and/or is expressed at the normal level;

(c) performing DNA fingerprinting to compare the restriction pattern produced when a restriction enzyme cuts nucleic acid in the sample with the restriction pattern obtained from the nucleotide sequence shown in Figure 7 or from a known mutant, allele or variant thereof;

(d) contacting the sample with a specific binding member capable of binding to nucleic acid including the nucleotide sequence as set out in Figure 7 or a fragment thereof, or a mutant, allele or variant thereof, the specific binding member including nucleic acid hybridisable with the sequence of Figure 7 or a polypeptide including a binding domain with specificity for nucleic acid including the sequence of Figure 7 or the polypeptide encoded by it, or a mutated form thereof, and

determining binding of the specific binding member;

(e) performing PCR involving one or more primers based on the nucleotide sequence shown in Figure 7 to screen the sample for nucleic acid including the nucleotide sequence of Figure 7 or a mutant, allele or variant thereof.

When screening for a resistance allele nucleic acid, the nucleic acid in the sample will initially be amplified, e.g. using PCR, to increase the amount of the analyte as compared to other sequences present in the sample. This allows the target sequences to be detected with a high degree of sensitivity if they are present in the sample. This initial step may be avoided by using highly sensitive array techniques that are becoming increasingly important in the art.

A variant form of the gene may contain one or more insertions, deletions, substitutions and/or additions of one or more nucleotides compared with the wild-type sequence (such as shown in Table 1) which may or may not disrupt the gene function. Differences at the nucleic acid level are not necessarily reflected by a difference in the amino acid sequence of the encoded polypeptide. However, a mutation or other difference in a gene may result in a frame-shift or stop codon, which could seriously affect the nature of the polypeptide produced (if any), or a point mutation or gross mutational change to the encoded polypeptide, including insertion, deletion, substitution and/or addition of one or more amino acids or regions in the polypeptide. A mutation in a promoter sequence or other regulatory region may prevent or



reduce expression from the gene or affect the processing or stability of the mRNA transcript.

Tests may be carried out on preparations containing genomic DNA, cDNA and/or mRNA. Testing cDNA or mRNA has the advantage of the complexity of the nucleic acid being reduced by the absence of intron sequences, but the possible disadvantage of extra time and effort being required in making the preparations. RNA is more difficult to manipulate than DNA because of the wide-spread occurrence of RN'ases.

Nucleic acid in a test sample may be sequenced and the sequence compared with the sequence shown in Figure 2, or other figure herein, to determine whether or not a difference is present. If so, the difference can be compared with known susceptibility alleles (e.g. as summarised in Table 1) to determine whether the test nucleic acid contains one or more of the variations indicated, or the difference can be investigated for association with disease resistance.

The amplified nucleic acid may then be sequenced as above, and/or tested in any other way to determine the presence or absence of a particular feature. Nucleic acid for testing may be prepared from nucleic acid removed from cells or in a library using a variety of other techniques such as restriction enzyme digest and electrophoresis.

Nucleic acid may be screened using a variant- or allele-specific probe. Such a probe corresponds in sequence to a region of the gene, or its complement, containing a sequence alteration known to be associated with disease resistance.

Under suitably stringent conditions, specific hybridisation of such a probe to test nucleic acid is indicative of the presence of the sequence alteration in the test nucleic acid. For efficient screening purposes, more than one probe may be used  
5 on the same test sample.

Allele- or variant-specific oligonucleotides may similarly be used in PCR to specifically amplify particular sequences if present in a test sample. Assessment of whether a PCR band contains a gene variant may be carried out in a number of ways  
10 familiar to those skilled in the art. The PCR product may for instance be treated in a way that enables one to display the mutation or polymorphism on a denaturing polyacrylamide DNA sequencing gel, with specific bands that are linked to the gene variants being selected.

15 An alternative or supplement to looking for the presence of variant sequences in a test sample is to look for the presence of the normal sequence, e.g. using a suitably specific oligonucleotide probe or primer.

Approaches which rely on hybridisation between a probe and  
20 test nucleic acid and subsequent detection of a mismatch may be employed. Under appropriate conditions (temperature, pH etc.), an oligonucleotide probe will hybridise with a sequence which is not entirely complementary. The degree of base-pairing between the two molecules will be sufficient for them to anneal  
25 despite a mis-match. Various approaches are well known in the art for detecting the presence of a mis-match between two annealing nucleic acid molecules.

For instance, RN'ase A cleaves at the site of a mis-match. Cleavage can be detected by electrophoresing test nucleic acid to which the relevant probe or probe has annealed and looking for smaller molecules (i.e. molecules with higher  
5 electrophoretic mobility) than the full length probe/test hybrid. Other approaches rely on the use of enzymes such as resolvases or endonucleases.

Thus, an oligonucleotide probe that has the sequence of a region of the normal gene (either sense or anti-sense strand)  
10 in which mutations associated with disease resistance are known to occur (e.g. see Table 1) may be annealed to test nucleic acid and the presence or absence of a mis-match determined. Detection of the presence of a mis-match may indicate the presence in the test nucleic acid of a mutation associated with  
15 disease resistance. On the other hand, an oligonucleotide probe that has the sequence of a region of the gene including a mutation associated with disease resistance may be annealed to test nucleic acid and the presence or absence of a mis-match determined. The presence of a mis-match may indicate that the  
20 nucleic acid in the test sample has the normal sequence, or a different mutant or allele sequence. In either case, a battery of probes to different regions of the gene may be employed.

The presence of differences in sequence of nucleic acid molecules may be detected by means of restriction enzyme  
25 digestion, such as in a method of DNA fingerprinting where the restriction pattern produced when one or more restriction enzymes are used to cut a sample of nucleic acid is compared

with the pattern obtained when a sample containing the normal gene or a variant or allele is digested with the same enzyme or enzymes.

The presence or absence of a lesion in a promoter or other  
5 regulatory sequence may also be assessed by determining the level of mRNA production by transcription or the level of polypeptide production by translation from the mRNA.

Nucleic acid isolated and/or purified from one or more cells of a plant or a nucleic acid library derived from nucleic  
10 acid isolated and/or purified from cells (e.g. a cDNA library derived from mRNA isolated from the cells), may be probed under conditions for selective hybridisation and/or subjected to a specific nucleic acid amplification reaction such as the polymerase chain reaction (PCR).

15 A method may include hybridisation of one or more (e.g. two) probes or primers to target nucleic acid. Where the nucleic acid is double-stranded DNA, hybridisation will generally be preceded by denaturation to produce single-stranded DNA. The hybridisation may be as part of a PCR  
20 procedure, or as part of a probing procedure not involving PCR. An example procedure would be a combination of PCR and low stringency hybridisation. A screening procedure, chosen from the many available to those skilled in the art, is used to identify successful hybridisation events and isolate hybridised  
25 nucleic acid.

Binding of a probe to target nucleic acid (e.g. DNA) may be measured using any of a variety of techniques at the

097223741300

disposal of those skilled in the art. For instance, probes may be radioactively, fluorescently or enzymatically labelled. Other methods not employing labelling of probe include examination of restriction fragment length polymorphisms, amplification using PCR, RNAase cleavage and allele specific oligonucleotide probing.

Probing may employ the standard Southern blotting technique. For instance DNA may be extracted from cells and digested with different restriction enzymes. Restriction fragments may then be separated by electrophoresis on an agarose gel, before denaturation and transfer to a nitrocellulose filter. Labelled probe may be hybridised to the DNA fragments on the filter and binding determined. DNA for probing may be prepared from RNA preparations from cells.

Preliminary experiments may be performed by hybridising under low stringency conditions various probes to Southern blots of DNA digested with restriction enzymes. Suitable conditions would be achieved when a large number of hybridising fragments were obtained while the background hybridisation was low. Using these conditions nucleic acid libraries, e.g. cDNA libraries representative of expressed sequences, may be searched.

As noted, those skilled in the art are well able to employ suitable conditions of the desired stringency for selective hybridisation, taking into account factors such as oligonucleotide length and base composition, temperature and so on.

In some preferred embodiments of diagnostic assays according to the present invention, oligonucleotides according to the present invention that are fragments of any of the sequences shown in Figure 2, or any allele associated with disease resistance, e.g. as identified in Table 1, are at least about 10 nucleotides in length, more preferably at least about 15 nucleotides in length, more preferably at least about 20 nucleotides in length, more preferably about 30 nucleotides in length. Such fragments themselves individually represent aspects of the present invention. Fragments and other oligonucleotides may be used as primers or probes as discussed but may also be generated (e.g. by PCR) in methods concerned with determining the presence in a test sample of a sequence indicative of disease resistance.

There are various methods for determining the presence or absence in a test sample of a particular polypeptide, such as the polypeptide with the amino acid sequence shown in Figure 2, or other figure herein, or an amino acid sequence mutant, variant or allele thereof (e.g. including an alteration shown in Table 1).

A sample may be tested for the presence of a binding partner for a specific binding member such as an antibody (or mixture of antibodies), specific for one or more particular variants of the polypeptide shown in Figure 2, e.g. see Table 1.

In such cases, the sample may be tested by being contacted

with a specific binding member such as an antibody under appropriate conditions for specific binding, before binding is determined, for instance using a reporter system as discussed. Where a panel of antibodies is used, different reporting labels  
5 may be employed for each antibody so that binding of each can be determined.

A specific binding member such as an antibody may be used to isolate and/or purify its binding partner polypeptide from a test sample, to allow for sequence and/or biochemical analysis  
10 of the polypeptide to determine whether it has the sequence and/or properties of the wild-type polypeptide or a particular mutant, variant or allele thereof. Amino acid sequence is routine in the art using automated sequencing machines.

15 The use of diagnostic tests for *mlo* alleles allows the researcher or plant breeder to establish, with full confidence and independent from time consuming resistance tests, whether or not a desired allele is present in the plant of interest (or a cell thereof), whether the plant is a representative of a  
20 collection of other genetically identical plants (e.g. an inbred variety or cultivar) or one individual in a sample of related (e.g. breeders' selection) or unrelated plants. The *mlo* alleles conferring the desirable disease resistance phenotype are recessive, and are not therefore detectable at  
25 the whole plant phenotype level when in a heterozygous condition in the presence of a wild-type *Mlo* allele. Phenotypic screening for the presence of such recessive alleles

is therefore only possible on material homozygous for the *mlo* locus and so delays substantially the generation in a plant breeding programme at which selection can be reliably and cost-effectively applied. In a backcross breeding programme where, for example, a breeder is aiming to introgress a desirable *mlo* allele into an elite adapted high performing target genotype, the *mlo* locus will be permanently in the heterozygous condition until selfing is carried out. Nucleic acid or polypeptide testing for the presence of the recessive allele avoids the need to test selfed progeny of backcross generation individuals, thus saving considerable time and money. In other types of breeding scheme based on selection and selfing of desirable individuals, nucleic acid or polypeptide diagnostics for the desirable *mlo* alleles in high throughput, low cost assays as provided by this invention, reliable selection for the desirable *mlo* alleles can be made at early generations and on more material than would otherwise be possible. This gain in reliability of selection plus the time saving by being able to test material earlier and without costly resistance phenotype screening is of considerable value in plant breeding.

By way of example for nucleic acid testing, the barley *mlo-5* resistance allele is characterized by a G- to A- nucleotide substitution in the predicted start codon of the *Mlo* gene (Table 1). The mutation may easily be detected by standard PCR amplification of a *Mlo* gene segment from genomic template DNA with the primers:

forward primer: 5'-GTTGCCACACTTTGCCACG-3'



reverse primer: 5'-AAGCCAAGACGACAATCAGA-3'

(for example), followed by digestion with the restriction enzyme *PshA1*. This generates a cleaved amplified polymorphic sequences (CAPS) marker which may be displayed using conventional agarose gel electrophoresis. Presence of a 769 bp fragment is indicative of the presence of the *mlo-5* allele.

The *mlo-9* resistance allele is characterized by a C- to T-nucleotide substitution (Table 1). This allele is of particular relevance since it is used frequently in breeding material. The mutational event may be easily detected using the primers:

forward primer 5'-GRRGCCACACTTTGCCACG-3'

reverse primer 5'-AAGCCAAGACGACAATCAGA-3'

(for example) and subsequent digestion of genomic amplification products with the restriction enzyme *Hha1*. This generates a CAPS marker which may be displayed by conventional agarose gel electrophoresis. The presence of a 374 bp fragment is indicative of the presence of *mlo-9*.

A third, particularly interesting allele is *mlo-12*, characterised by a substitution at residue 240, specifically a Phe240 to leucine replacement. This may result from a C720 to A substitution in the encoding nucleotide sequence (Table 1). This is the only currently documented *mlo* allele for which conclusive evidence is available that the altered protein retains residual wild-type activity (Hentrich, 1979, Arch. Züchtungsvorsch., Berlin 9, S. 283-291). *mlo-12* exhibits no detectable spontaneous cell death reaction but confers a

32

sufficient level of resistance to pathogens such as the powdery mildew fungus. *mlo-12* may therefore be the allele of choice in breeding programs if minimal pleiotropic effects (spontaneous cell death) are desirable after introgression of the *mlo* resistance in elite breeding lines. Furthermore, the molecular site of the amino acid substitution within the Mlo protein allows the design of alleles with a residual wild-type activity, and also the obtention of interacting and/or inhibitory molecules, reducing undesirable pleiotropic effects from a complete loss of function of the Mlo protein.

Nucleic acid-based determination of the presence or absence of *mlo* alleles may be combined with determination of the genotype of the flanking linked genomic DNA and other unlinked genomic DNA using established sets of markers such as RFLPs, microsatellites or SSRs, AFLPs, RAPDs etc. This enables the researcher or plant breeder to select for not only the presence of the desirable *mlo* allele but also for individual plant or families of plants which have the most desirable combinations of linked and unlinked genetic background. Such recombinations of desirable material may occur only rarely within a given segregating breeding population or backcross progeny. Direct assay of the *mlo* locus as afforded by the present invention allows the researcher to make a stepwise approach to fixing (making homozygous) the desired combination of flanking markers and *mlo* alleles, by first identifying individuals fixed for one flanking marker and then identifying progeny fixed on the other side of the *mlo* locus all the time

knowing with confidence that the desirable *mlo* allele is still present.

The present disclosure provides sufficient information for a person skilled in the art to obtain genomic DNA sequence for any given new or existing *mlo* allele and devise a suitable nucleic acid- and/or polypeptide-based diagnostic assay.

Existing *mlo* alleles to which this may be applied include, for example, *mlo*-1, *mlo*-3, *mlo*-4, *mlo*-5, *mlo*-6, *mlo*-7, *mlo*-8, *mlo*-9, *mlo*-10, *mlo*-12, *mlo*-13, *mlo*-16, *mlo*-17, *mlo*-26 and *mlo*-28, for all of which sequence information is provided herein (see e.g. Figure 2 and Table 1). In designing a nucleic acid assay account is taken of the distinctive variation in sequence that characterises the particular variant allele. Thus, the present invention extends to an oligonucleotide fragment of a *mlo* allele, having a sequence which allows it to hybridise specifically to that allele as compared with other *mlo* alleles. Such an oligonucleotide spans a nucleotide at which a *mlo* mutation occurs, and may include the mutated nucleotide at or towards its 3' or 5' end. Such an oligonucleotide may hybridise with the sense or anti-sense strand. The variation may be within the coding sequence of the *mlo* gene, or may lie within an intron sequence or in an upstream or downstream non-coding sequence, wherein disruption affects or is otherwise related to the lesion in *Mlo* that results in the mildew resistant phenotype.

The *mlo*-9 allele is widely but not exclusively used in plant breeding (J Helms Jorgensen - *Euphytica* (1992) 63: 141-

152), *mlo-11* is also used. Use of *mlo* mutants in practical breeding has largely been restricted to spring barley, because the spontaneous cell death response associated with many of the mutant alleles appears to represent a penalty to plant growth and performance when incorporated into high yielding winter barley genotypes. However different *mlo* alleles have different degrees of associated spontaneous cell death response, and thus some, either existing or newly created from mutagenesis programmes or isolated as spontaneous mutants, are more suitable than others for incorporation into winter barley backgrounds. The *mlo-12* allele may be particularly suitable since no detectable pleiotropic effects occur despite conferring a sufficient level of pathogen resistance. The use of *mlo* based mildew resistance more widely in winter barleys will have significant value for barley growers as well as significant economic and environmental implications such as reduced use of fungicide inputs with their associated treatment costs. The provision of nucleic acid diagnostics as provided herein enables rapid and accurate deployment of new and existing *mlo* alleles into winter barley germplasm.

Plants which include a plant cell according to the invention are also provided, along with any part or propagule thereof, seed, selfed or hybrid progeny and descendants. A plant according to the present invention may be one which does not breed true in one or more properties. Plant varieties may be excluded, particularly registrable plant varieties according

to Plant Breeders' Rights. It is noted that a plant need not be considered a "plant variety" simply because it contains stably within its genome a transgene, introduced into a cell of the plant or an ancestor thereof.

5 In addition to a plant, the present invention provides any clone of such a plant, seed, selfed or hybrid progeny and descendants, and any part of any of these, such as cuttings, seed. The invention provides any plant propagule, that is any part which may be used in reproduction or propagation, sexual  
10 or asexual, including cuttings, seed and so on. Also encompassed by the invention is a plant which is a sexually or asexually propagated off-spring, clone or descendant of such a plant, or any part or propagule of said plant, off-spring, clone or descendant.

15 A further aspect of the present invention provides a method of making a plant cell involving introduction of the sequence (e.g. as part of a suitable vector) into a plant cell and causing or allowing recombination between the vector and the plant cell genome to introduce the sequence of nucleotides  
20 into the genome.

Following transformation of a plant cell a plant may be regenerated.

25 The invention further provides a method of modulating *Mlo* expression in a plant, which may modulate a defence response in the plant, comprising expression of a heterologous *Mlo* gene sequence (or mutant, allele, variant or homologue thereof, as

20

15

20

25

or under control of an inducible promoter. Furthermore, mutants, variants and derivatives of the wild-type gene, e.g. with higher or lower activity than wild-type, may be used in

place of the endogenous gene. Nucleic acid heterologous, or exogenous or foreign, to a plant cell may be non-naturally occurring in cells of that type, variety or species. Thus, nucleic acid may include a coding sequence of or derived from a particular type of plant cell or species or variety of plant, placed within the context of a plant cell of a different type or species or variety of plant. A further possibility is for a nucleic acid sequence to be placed within a cell in which it or a homologue is found naturally, but wherein the nucleic acid sequence is linked and/or adjacent to nucleic acid which does not occur naturally within the cell, or cells of that type or species or variety of plant, such as operably linked to one or more regulatory sequences, such as a promoter sequence, for control of expression. A sequence within a plant or other host cell may be identifiably heterologous, exogenous or foreign.

Down-regulation of wild-type *Mlo* gene function leads to stimulation of a constitutive defence response. This may be achieved in a number of different ways, as illustrated below.

The nucleic acid according to the invention may be placed under the control of an inducible gene promoter thus placing expression under the control of the user.

In a further aspect the present invention provides a gene construct comprising an inducible promoter operatively linked to a nucleotide sequence provided by the present invention. As discussed, this enables control of expression of the gene. The invention also provides plants transformed with said gene construct and methods comprising introduction of such a

construct into a plant cell and/or induction of expression of a construct within a plant cell, e.g. by application of a suitable stimulus, such as an effective exogenous inducer or endogenous signal.

5       The term "inducible" as applied to a promoter is well understood by those skilled in the art. In essence, expression under the control of an inducible promoter is "switched on" or increased in response to an applied stimulus (which may be generated within a cell or provided exogenously). The nature of  
10   the stimulus varies between promoters. Some inducible promoters cause little or undetectable levels of expression (or no expression) in the absence of the appropriate stimulus. Other inducible promoters cause detectable constitutive expression in the absence of the stimulus. Whatever the level of expression  
15   is in the absence of the stimulus, expression from any inducible promoter is increased in the presence of the correct stimulus. The preferable situation is where the level of expression increases upon application of the relevant stimulus by an amount effective to alter a phenotypic characteristic.  
20   Thus an inducible (or "switchable") promoter may be used which causes a basic level of expression in the absence of the stimulus which level is too low to bring about a desired phenotype (and may in fact be zero). Upon application of the stimulus, expression is increased (or switched on) to a level  
25   which brings about the desired phenotype.

Suitable promoters include the Cauliflower Mosaic Virus 35S (CaMV 35S) gene promoter that is expressed at a high level

097223977-112800



in virtually all plant tissues (Benfey et al, (1990a) EMBO J 9: 1677-1684); the cauliflower meri 5 promoter that is expressed in the vegetative apical meristem as well as several well localised positions in the plant body, eg inner phloem, flower primordia, branching points in root and shoot (Medford, J.I. (1992) *Plant Cell* 4, 1029-1039; Medford et al, (1991) *Plant Cell* 3, 359-370) and the *Arabidopsis thaliana* LEAFY promoter that is expressed very early in flower development (Weigel et al, (1992) *Cell* 69, 843-859).

10 An aspect of the present invention is the use of nucleic acid according to the invention in the production of a transgenic plant.

When introducing a chosen gene construct into a cell, certain considerations must be taken into account, well known to those skilled in the art. The nucleic acid to be inserted should be assembled within a construct which contains effective regulatory elements which will drive transcription. There must be available a method of transporting the construct into the cell. Once the construct is within the cell membrane, 15 integration into the endogenous chromosomal material either will or will not occur. Finally, as far as plants are concerned the target cell type must be such that cells can be regenerated into whole plants. 20

Plants transformed with the DNA segment containing the 25 sequence may be produced by standard techniques which are already known for the genetic manipulation of plants. DNA can be transformed into plant cells using any suitable technology,

such as a disarmed Ti-plasmid vector carried by *Agrobacterium* exploiting its natural gene transfer ability (EP-A-270355, EP-A-0116718, NAR 12(22) 8711 - 87215 1984), particle or microprojectile bombardment (US 5100792, EP-A-444882, EP-A-434616) microinjection (WO 92/09696, WO 94/00583, EP 331083, EP 175966, Green et al. (1987) *Plant Tissue and Cell Culture*, Academic Press), electroporation (EP 290395, WO 8706614) other forms of direct DNA uptake (DE 4005152, WO 9012096, US 4684611), liposome mediated DNA uptake (e.g. Freeman et al. *Plant Cell Physiol.* 29: 1353 (1984)), or the vortexing method (e.g. Kindle, *PNAS U.S.A.* 87: 1228 (1990d) Physical methods for the transformation of plant cells are reviewed in Oard, 1991, *Biotech. Adv.* 9: 1-11.

*Agrobacterium* transformation is widely used by those skilled in the art to transform dicotyledonous species. Recently, there has been substantial progress towards the routine production of stable, fertile transgenic plants in almost all economically relevant monocot plants (Toriyama, et al. (1988) *Bio/Technology* 6, 1072-1074; Zhang, et al. (1988) *Plant Cell Rep.* 7, 379-384; Zhang, et al. (1988) *Theor Appl Genet* 76, 835-840; Shimamoto, et al. (1989) *Nature* 338, 274-276; Datta, et al. (1990) *Bio/Technology* 8, 736-740; Christou, et al. (1991) *Bio/Technology* 9, 957-962; Peng, et al. (1991) International Rice Research Institute, Manila, Philippines 563-574; Cao, et al. (1992) *Plant Cell Rep.* 11, 585-591; Li, et al. (1993) *Plant Cell Rep.* 12, 250-255; Rathore, et al. (1993) *Plant Molecular Biology* 21, 871-884; Fromm, et al. (1990)

008217 4422/60

*Bio/Technology* 8, 833-839; Gordon-Kamm, et al. (1990) *Plant Cell* 2, 603-618; D'Halluin, et al. (1992) *Plant Cell* 4, 1495-1505; Walters, et al. (1992) *Plant Molecular Biology* 18, 189-200; Koziel, et al. (1993) *Biotechnology* 11, 194-200; Vasil, I. K. (1994) *Plant Molecular Biology* 25, 925-937; Weeks, et al. (1993) *Plant Physiology* 102, 1077-1084; Somers, et al. (1992) *Bio/Technology* 10, 1589-1594; WO92/14828). In particular, *Agrobacterium* mediated transformation is now emerging also as an highly efficient alternative transformation method in monocots (Hiei et al. (1994) *The Plant Journal* 6, 271-282).

The generation of fertile transgenic plants has been achieved in the cereals rice, maize, wheat, oat, and barley (reviewed in Shimamoto, K. (1994) *Current Opinion in Biotechnology* 5, 158-162.; Vasil, et al. (1992) *Bio/Technology* 10, 667-674; Vain et al., 1995, *Biotechnology Advances* 13 (4): 653-671; Vasil, 1996, *Nature Biotechnology* 14 page 702).

Microprojectile bombardment, electroporation and direct DNA uptake are preferred where *Agrobacterium* is inefficient or ineffective. Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process, eg bombardment with *Agrobacterium* coated microparticles (EP-A-486234) or microprojectile bombardment to induce wounding followed by co-cultivation with *Agrobacterium* (EP-A-486233).

Following transformation, a plant may be regenerated, e.g. from single cells, callus tissue or leaf discs, as is standard in the art. Almost any plant can be entirely regenerated from

cells, tissues and organs of the plant. Available techniques are reviewed in Vasil et al., *Cell Culture and Somatic Cell Genetics of Plants*, Vol I, II and III, *Laboratory Procedures and Their Applications*, Academic Press, 1984, and Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989.

The particular choice of a transformation technology will be determined by its efficiency to transform certain plant species as well as the experience and preference of the person practising the invention with a particular methodology of choice. It will be apparent to the skilled person that the particular choice of a transformation system to introduce nucleic acid into plant cells is not essential to or a limitation of the invention, nor is the choice of technique for plant regeneration.

In the present invention, expression may be achieved by introduction of the nucleotide sequence in a sense orientation. Thus, the present invention provides a method of modulation of a defence response in a plant, the method comprising causing or allowing expression of nucleic acid according to the invention within cells of the plant. Generally, it will be desirable to stimulate the defence response, and this may be achieved by disrupting *Mlo* gene function.

Down-regulation of expression of a target gene may be achieved using anti-sense technology or "sense regulation" ("co-suppression").

In using anti-sense genes or partial gene sequences to

09722377.43800

down-regulate gene expression, a nucleotide sequence is placed under the control of a promoter in a "reverse orientation" such that transcription yields RNA which is complementary to normal mRNA transcribed from the "sense" strand of the target gene.

5 See, for example, Rothstein et al, 1987; Smith et al, (1988) *Nature* 334, 724-726; Zhang et al, (1992) *The Plant Cell* 4, 1575-1588, English et al., (1996) *The Plant Cell* 8, 179-188. Antisense technology is also reviewed in Bourque, (1995), *Plant Science* 105, 125-149, and Flavell, (1994) *PNAS USA* 91, 3490-10 3496.

An alternative is to use a copy of all or part of the target gene inserted in sense, that is the same, orientation as the target gene, to achieve reduction in expression of the target gene by co-suppression. See, for example, van der Krol 15 et al., (1990) *The Plant Cell* 2, 291-299; Napoli et al., (1990) *The Plant Cell* 2, 279-289; Zhang et al., (1992) *The Plant Cell* 4, 1575-1588, and US-A-5,231,020.

The complete sequence corresponding to the coding sequence (in reverse orientation for anti-sense) need not be used. For 20 example fragments of sufficient length may be used. It is a routine matter for the person skilled in the art to screen fragments of various sizes and from various parts of the coding sequence to optimise the level of anti-sense inhibition. It may be advantageous to include the initiating methionine ATG 25 codon, and perhaps one or more nucleotides upstream of the initiating codon. A further possibility is to target a conserved sequence of a gene, e.g. a sequence that is

characteristic of one or more genes, such as a regulatory sequence. Antisense constructs may involve 3' end or 5' end sequences of *Mlo* or homologues. In cases where several *Mlo* homologues exist in a plant species, the involvement of 5'- and 3'-end untranslated sequences in the construct will enhance specificity of silencing.

The sequence employed may be about 500 nucleotides or less, possibly about 400 nucleotides, about 300 nucleotides, about 200 nucleotides, or about 100 nucleotides. It may be possible to use oligonucleotides of much shorter lengths, 14-23 nucleotides, although longer fragments, and generally even longer than about 500 nucleotides are preferable where possible, such as longer than about 600 nucleotides, than about 700 nucleotides, than about 800 nucleotides, than about 1000 nucleotides, than about 1200 nucleotides, than about 1400 nucleotides, or more.

It may be preferable that there is complete sequence identity in the sequence used for down-regulation of expression of a target sequence, and the target sequence, though total complementarity or similarity of sequence is not essential. One or more nucleotides may differ in the sequence used from the target gene. Thus, a sequence employed in a down-regulation of gene expression in accordance with the present invention may be a wild-type sequence (e.g. gene) selected from those available, or a mutant, derivative, variant or allele, by way of insertion, addition, deletion or substitution of one or more nucleotides, of such a sequence. The sequence need not

include an open reading frame or specify an RNA that would be translatable. It may be preferred for there to be sufficient homology for the respective anti-sense and sense RNA molecules to hybridise. There may be down regulation of gene expression even where there is about 5%, 10%, 15% or 20% or more mismatch between the sequence used and the target gene.

Generally, the transcribed nucleic acid may represent a fragment of an Mlo gene, such as including a nucleotide sequence shown in Figure 2, or the complement thereof, or may be a mutant, derivative, variant or allele thereof, in similar terms as discussed above in relation to alterations being made to a coding sequence and the homology of the altered sequence. The homology may be sufficient for the transcribed anti-sense RNA to hybridise with nucleic acid within cells of the plant, though irrespective of whether hybridisation takes place the desired effect is down-regulation of gene expression.

Anti-sense regulation may itself be regulated by employing an inducible promoter in an appropriate construct.

Constructs may be expressed using the natural promoter, by a constitutively expressed promoter such as the CaMV 35S promoter, by a tissue-specific or cell-type specific promoter, or by a promoter that can be activated by an external signal or agent. The CaMV 35S promoter but also the rice *actin1* and maize ubiquitin promoters have been shown to give high levels of reporter gene expression in rice (Fujimoto et al., (1993) *Bio/Technology* 11, 1151-1155; Zhang, et al., (1991) *Plant Cell* 3, 1155-1165; Cornejo et al., (1993) *Plant Molecular Biology*

23, 567-581).

For use in anti-sense regulation, nucleic acid including a nucleotide sequence complementary to a coding sequence of a Mlo gene (i.e. including homologues), or a fragment of a said coding sequence suitable for use in anti-sense regulation of expression, is provided. This may be DNA and under control of an appropriate regulatory sequence for anti-sense transcription in cells of interest.

Thus, the present invention also provides a method of conferring pathogen resistance on a plant, the method including causing or allowing anti-sense transcription from heterologous nucleic acid according to the invention within cells of the plant.

The present invention further provides the use of the nucleotide sequence of Figure 2 or a fragment, mutant, derivative, allele, variant or homologue thereof, such as any sequence shown or identified herein, for down-regulation of gene expression, particularly down-regulation of expression of an Mlo gene or homologue thereof, preferably in order to confer pathogen resistance on a plant.

When additional copies of the target gene are inserted in sense, that is the same, orientation as the target gene, a range of phenotypes is produced which includes individuals where over-expression occurs and some where under-expression of protein from the target gene occurs. When the inserted gene is only part of the endogenous gene the number of under-expressing individuals in the transgenic population increases. The



mechanism by which sense regulation occurs, particularly down-regulation, is not well-understood. However, this technique is well-reported in scientific and patent literature and is used routinely for gene control. See, for example, van  
5 der Krol et al., (1990) *The Plant Cell* 2, 291-229; Napoli et al., (1990) *The Plant Cell* 2, 279-289; Zhang et al, 1992 *The Plant Cell* 4, 1575-1588.

Again, fragments, mutants and so on may be used in similar terms as described above for use in anti-sense regulation.

10 Thus, the present invention also provides a method of conferring pathogen resistance on a plant, the method including causing or allowing expression from nucleic acid according to the invention within cells of the plant. This may be used to suppress Mlo activity. Here the activity of the product is  
15 preferably suppressed as a result of under-expression within the plant cells.

As noted, Mlo down-regulation may promote activation of a defence response, which may in turn confer or augment pathogen resistance of the plant, especially resistance to powdery  
20 mildew and/or rust (e.g. yellow rust).

Thus, the present invention also provides a method of modulating Mlo function in a plant, the method comprising causing or allowing expression from nucleic acid according to the invention within cells of the plant to suppress endogenous  
25 Mlo expression.

Modified versions of Mlo may be used to down-regulate endogenous Mlo function. For example mutants, variants,

derivatives etc., may be employed. For instance, expression of a *mlo* mutant sequence at a high level may out-compete activity of endogenous *Mlo*.

Reduction of *Mlo* wild type activity may be achieved by  
5 using ribozymes, such as replication ribozymes, e.g. of the hammerhead class (Haseloff and Gerlach, 1988, *Nature* 334: 585-591; Feyter et al. *Mol.*, 1996, *Gen. Genet.* 250: 329-338).

Another way to reduce *Mlo* function in a plant employs  
transposon mutagenesis (reviewed by Osborne et al., (1995)  
10 *Current Opinion in Cell Biology* 7, 406-413). Inactivation of genes has been demonstrated via a 'targeted tagging' approach using either endogenous mobile elements or heterologous cloned transposons which retain their mobility in alien genomes. *Mlo* alleles carrying any insertion of known sequence could be  
15 identified by using PCR primers with binding specificities both in the insertion sequence and the *Mlo* homologue. 'Two-element systems' could be used to stabilize the transposon within inactivated alleles. In the two-element approach, a T-DNA is constructed bearing a non-autonomous transposon containing  
20 selectable or screenable marker gene inserted into an excision marker. Plants bearing these T-DNAs are crossed to plants bearing a second T-DNA expressing transposase function. Hybrids are double-selected for excision and for the marker within the transposon yielding  $F_2$  plants with transposed elements. The  
25 two-element approach has a particular advantage with respect to *Ac/Ds* of maize, as the transposed *Ds* is likely to be unlinked to the transposase, facilitating outcrossing and stabilization

of the *Ds* insertion (Jones et al., (1994) *Science* 266, 789-793; Osborne et al., (1995) *Current Opinion in Cell Biology* 7, 406-413).

5       The *mlo*-based powdery mildew resistance is caused by the inactivation of the *Mlo* wild type allele, resulting in a recessive resistance phenotype. Substances that inhibit the activity of the *Mlo* wild type protein may be used to induce the resistance phenotype.

10       An important hint that complete inactivation of *Mlo* expression is not essential and may even be detrimental is provided by the description of mutagen-induced *mlo* resistance alleles that are likely to have retained residual wild type allele activity. These alleles exhibit no detectable  
15       spontaneous leaf necrosis which negatively affects photosynthesis rates and yield (Hentrich, W (1979) *Arch. Züchtungsvorsch., Berlin* 9, S. 283-291).

20       The *Mlo* protein is predicted to be membrane-anchored by seven transmembrane helices (see e.g. Figure 7). This structure prediction has been reinforced by recent analysis of *Mlo* homologues in rice and *Arabidopsis thaliana*. Structure prediction of the *Arabidopsis thaliana* homologue also suggests the presence of seven transmembrane helices. A comparison of the *Mlo* homologues revealed in addition conserved cysteine  
25       residues in the putative extracellular loops 1 and 3 and high probabilities of amphipathic helices in the second intracellular loop adjacent to the predicted transmembrane

helices 3 and 4. These conserved structural motifs in the family of Mlo proteins are reminiscent of G protein coupled receptors (GPCR) described extensively in mammalian systems. GPCRs are known to be activated by ligands and to amplify signals intracellularly via heterotrimeric G proteins. Without in any way providing a limitation on the nature or scope of any aspect of the present invention, it is predicted that Mlo activates an inhibitory G alpha subunit of heterotrimeric G proteins, thus leading to a downregulation of as yet unknown effector proteins.

The provision herein of Mlo sequence information enables the identification of antagonists of function of the Mlo protein (e.g. GPCR function). Antagonists of Mlo may block receptor activation by its unknown genuine ligand, mimicking recessive mutations in the Mlo gene. Such Mlo antagonists may be used as crop protection compounds, for example applied externally to the plant or crop or, where the compound is peptidyl in nature, delivered internally via a biological vector (e.g. recombinant infecting viral particle expressing the antagonistic molecule within target plant cells) or via a transgenic route (plants or plant cells genetically modified to express the antagonist molecule, perhaps under control of a promoter inducible by an externally applied compound (eg GST-II promoter from maize - Jepson et al Plant Molecular Biology 26:1855-1866 (1994)) allowing control over the timing of expression of the mlo inactivation phenotype.

Leaf segments of Mlo wild type plants may be tested with a

0972377 41300

test substance, e.g. from a random or combinatorial compound library, for resistance upon challenge with pathogen such as powdery mildew. The detached leaf segment assay is used as a standard test system to score for susceptibility/resistance upon inoculation with powdery mildew spores. Leaf segments of 7-day-old seedlings of the genotype *Mlo RorI* may be placed on agar, for example individual wells of 96-well microtiter plates containing 50  $\mu$ l agar. Different compounds may be applied to the agar surface in each well at a concentration of about 1ppm dissolved in DMSO. Around seven days after inoculation of the detached leaf segments with pathogen, such as spores of a virulent powdery mildew isolate, compounds which induce resistance may be recognised by the absence of fungal mycelium on leaf segments in the microtiter plates.

A further selection may be used to discriminate between compounds that act in the *mlo* pathway and those that confer resistance by other mechanisms, or those which exhibit a direct fungitoxic activity. For this purpose mutants in genes (*Ror* genes) which may be required for *mlo* resistance (Freialdenhoven et al., (1996), The Plant Cell 8, 5-14) may be used. Mutants of these genes confer susceptibility to powdery mildew attack despite the presence of *mlo* resistance alleles. Plants of the genotype *Mlo rorl* (wild type *Mlo* protein and defective *Rorl* gene) may be used, for example, to test compounds which induce resistance on *Mlo Rorl* genotypes but exhibit susceptibility on the *Mlo rorl* genotype, enabling selection of candidate *Mlo* antagonists. Testing candidate compounds identified using a

leaf segment test may be used to drastically reduce the number of candidate compounds for further *in vitro* tests.

A further selection step of candidate antagonists may involve heterologous expression of the Mlo protein or a  
5 fragment thereof (e.g. in a baculovirus insect cell system) and subsequent binding assays with labelled molecules. Specific binding of compounds to cell lines expressing wild type Mlo protein is a good indicator of their antagonistic mode of action. Analysis of the deduced Mlo protein sequence has  
10 provided strong evidence that the protein is anchored in the membrane via seven transmembrane helices and may represent a novel member of the so-called serpentine receptor family. The conclusion is supported by the sequence data derived from homologous genes identified in barley, rice and *Arabidopsis*.  
15 Seven transmembrane proteins have been shown to be expressed at high level in the Baculovirus/insect cell system (up to  $10^7$  molecules per cell - Tate and Grisshamer, 1996, *TIBTECH* 14: 426-430). Since the family of Mlo proteins appears to be restricted to the plant kingdom, this provides a low-background  
20 environment for compound tests. Candidate compounds which are labelled, radioactively or non-radioactively, may be tested for specific binding to Sf9 insect cells expressing the Mlo protein after infection with a recombinant baculovirus construct. Specificity of the binding may be tested further by Sf9  
25 expression of mutant *mlo* proteins which carry characterised mutations (e.g. as in Table 1) leading *in vivo* to resistance.

Thus, in various further aspects the present invention relates to assays for substances able to interfere with Mlo function, i.e. confer a *mlo* mutant phenotype, such substances themselves and uses thereof.

5       The use of Mlo in identifying and/or obtaining a substance which inhibits Mlo function is further provided by the present invention, as is the use of Mlo in identifying and/or obtaining a substance which induces pathogen resistance in a plant.

10       Agents useful in accordance with the present invention may be identified by screening techniques which involve determining whether an agent under test inhibits or disrupts Mlo function to induce an *mlo* phenotype. Candidate inhibitors are substances which bind Mlo.

15       It should of course be noted that references to "Mlo" in relation to assays and screens should be taken to refer to homologues, such as in other species, including rice and wheat, not just in barley, also appropriate fragments, variants, alleles and derivatives thereof. Assessment of whether a test  
20       substance is able to bind the Mlo protein does not necessarily require the use of full-length Mlo protein. A suitable fragment may be used (or a suitable analogue or variant thereof).

25       Suitable fragments of Mlo include those which include residues known to be crucial for Mlo function as identified by *mlo* mutant alleles (Table 1). Smaller fragments, and analogues and variants of this fragment may similarly be employed, e.g.

as identified using techniques such as deletion analysis or alanine scanning.

Furthermore, one class of agents that can be used to disrupt Mlo activity are peptides fragments of it. Such peptides tend to be short, and may be about 40 amino acids in length or less, preferably about 35 amino acids in length or less, more preferably about 30 amino acids in length, or less, more preferably about 25 amino acids or less, more preferably about 20 amino acids or less, more preferably about 15 amino acids or less, more preferably about 10 amino acids or less, or 9, 8, 7, 6, 5 or less in length. The present invention also encompasses peptides which are sequence variants or derivatives of a wild type Mlo sequence, but which retain ability to interfere with Mlo function, e.g. to induce an mlo mutant phenotype. Where one or more additional amino acids are included, such amino acids may be from Mlo or may be heterologous or foreign to Mlo. A peptide may also be included within a larger fusion protein, particularly where the peptide is fused to a non-Mlo (i.e. heterologous or foreign) sequence, such as a polypeptide or protein domain.

Peptides may be generated wholly or partly by chemical synthesis. The compounds of the present invention can be readily prepared according to well-established, standard liquid or, preferably, solid-phase peptide synthesis methods, general descriptions of which are broadly available (see, for example, in J.M. Stewart and J.D. Young, Solid Phase Peptide Synthesis, 2nd edition, Pierce Chemical Company, Rockford, Illinois

0972277 412800



1984), in M. Bodanzsky and A. Bodanzsky, The Practice of Peptide Synthesis, Springer Verlag, New York (1984); and Applied Biosystems 430A Users Manual, ABI Inc., Foster City, California), or they may be prepared in solution, by the liquid phase method or by any combination of solid-phase, liquid phase and solution chemistry, e.g. by first completing the respective peptide portion and then, if desired and appropriate, after removal of any protecting groups being present, by introduction of the residue X by reaction of the respective carbonic or sulfonic acid or a reactive derivative thereof.

Another convenient way of producing a peptidyl molecule according to the present invention (peptide or polypeptide) is to express nucleic acid encoding it, by use of nucleic acid in an expression system, as discussed elsewhere herein. This allows for peptide agents to be delivered to plants transgenically, by means of encoding nucleic acid. If coupled to an inducible promoter for expression under control of the user, this allows for flexibility in induction of an mlo phenotype and pathogen resistance. This may allow for any side-effects arising from interference with Mlo function to be moderated.

In one general aspect the present invention provides an assay method for a substance able to interact with the relevant region of Mlo, the method including:

(a) bringing into contact a Mlo polypeptide or peptide fragment thereof, or a variant, derivative or analogue

56

thereof, and a test compound; and

(b) determining interaction or binding between said polypeptide or peptide and the test compound.

A test compound found to interact with the relevant  
5 portion of Mlo may be tested for ability to modulate, e.g. disrupt or interfere with, Mlo function, as discussed already above.

Another general aspect of the present invention provides  
10 an assay method for a substance able to induce an mlo mutant phenotype in a plant, the method including:

(a) bringing into contact a plant or part thereof (e.g. leaf or leaf segment) and a test compound; and

(b) determining Mlo function and/or pathogen resistance  
15 and/or stimulation of a defence response in the plant.

Susceptibility or resistance to a pathogen may be determined by assessing pathogen growth, e.g. for powdery mildew the presence or absence, or extent, of mycelial growth.

Binding of a test compound to a polypeptide or peptide may  
20 be assessed in addition to ability of the test compound to stimulate a defence response in a plant. Such tests may be run in parallel or one test may be performed on a substance which tests positive in another test.

25 Of course, the person skilled in the art will design any appropriate control experiments with which to compare results obtained in test assays.

09/23/97 11:23:00

Performance of an assay method according to the present invention may be followed by isolation and/or manufacture and/or use of a compound, substance or molecule which tests positive for ability to modulate Mlo function and/or induce pathogen resistance, such as resistance to powdery mildew.

The precise format of an assay of the invention may be varied by those of skill in the art using routine skill and knowledge. For example, interaction between substances may be studied *in vitro* by labelling one with a detectable label and bringing it into contact with the other which has been immobilised on a solid support. Suitable detectable labels, especially for peptidyl substances include  $^{35}\text{S}$ -methionine which may be incorporated into recombinantly produced peptides and polypeptides. Recombinantly produced peptides and polypeptides may also be expressed as a fusion protein containing an epitope which can be labelled with an antibody.

An assay according to the present invention may also take the form of an *in vivo* assay. The *in vivo* assay may be performed in a cell line such as a yeast strain or mammalian cell line in which the relevant polypeptides or peptides are expressed from one or more vectors introduced into the cell.

For example, a polypeptide or peptide containing a fragment of Mlo or a peptidyl analogue or variant thereof as disclosed, may be fused to a DNA binding domain such as that of the yeast transcription factor GAL 4. The GAL 4 transcription factor includes two functional domains. These domains are the

0972237.43800

DNA binding domain (GAL4DBD) and the GAL4 transcriptional activation domain (GAL4TAD). By fusing such a polypeptide or peptide to one of those domains and another polypeptide or peptide to the respective counterpart, a functional GAL 4 transcription factor is restored only when two polypeptides or peptides of interest interact. Thus, interaction of the polypeptides or peptides may be measured by the use of a reporter gene probably linked to a GAL 4 DNA binding site which is capable of activating transcription of said reporter gene.

10 This assay format is described by Fields and Song, 1989, Nature 340; 245-246. This type of assay format can be used in both mammalian cells and in yeast. Other combinations of DNA binding domain and transcriptional activation domain are available in the art and may be preferred, such as the LexA DNA

15 binding domain and the VP60 transcriptional activation domain.

When looking for peptides or other substances which interact with Mlo, the Mlo polypeptide or peptide may be employed as a fusion with (e.g.) the LexA DNA binding domain, with test polypeptide or peptide (e.g. a random or

20 combinatorial peptide library) as a fusion with (e.g.) VP60. An increase in reporter gene expression (e.g. in the case of  $\beta$ -galactosidase a strengthening of the blue colour) results from the presence of a peptide which interacts with Mlo, which interaction is required for transcriptional activation of the

25  $\beta$ -galactosidase gene.

The amount of test substance or compound which may be

added to an assay of the invention will normally be determined by trial and error depending upon the type of compound used. Typically, from about 0.001 nM to 1mM or more concentrations of putative inhibitor compound may be used, for example from 0.01 nM to 100  $\mu$ M, e.g. 0.1 to 50  $\mu$ M, such as about 10  $\mu$ M. Greater concentrations may be used when a peptide is the test substance. Even a molecule which has a weak effect may be a useful lead compound for further investigation and development.

Compounds which may be used may be natural or synthetic chemical compounds used in drug screening programmes. Extracts of plants which contain several characterised or uncharacterised components may also be used. Antibodies directed to M10 or a fragment thereof form a further class of putative inhibitor compounds. Candidate inhibitor antibodies may be characterised and their binding regions determined to provide single chain antibodies and fragments thereof which are responsible for disrupting the interaction. Other candidate inhibitor compounds may be based on modelling the 3-dimensional structure of a polypeptide or peptide fragment and using rational drug design to provide potential inhibitor compounds with particular molecular shape, size and charge characteristics. It is worth noting, however, that combinatorial library technology provides an efficient way of testing a potentially vast number of different substances for ability to interact with and/or modulate the activity of a polypeptide. Such libraries and their use are known in the art, for all manner of natural products, small molecules and

peptides, among others. The use of peptide libraries may be preferred in certain circumstances.

Following identification of a substance or agent which  
5 modulates or affects Mlo function, the substance or agent may be investigated further. Furthermore, it may be manufactured and/or used in preparation, i.e. manufacture or formulation, of a composition for inducing pathogen resistance in a plant. These may be applied to plants, e.g. for inducing pathogen  
10 resistance, such as resistance to powdery mildew. A further aspect of the present invention provides a method of inducing pathogen resistance in a plant, the method including applying such a substance to the plant. A peptidyl molecule may be applied to a plant transgenically, by expression from encoding  
15 nucleic acid, as noted.

A polypeptide, peptide or other substance able to modulate or interfere with Mlo function, inducing pathogen resistance in a plant as disclosed herein, or a nucleic acid molecule  
20 encoding a peptidyl such molecule, may be provided in a kit, e.g. sealed in a suitable container which protects its contents from the external environment. Such a kit may include instructions for use.

25 Further aspects and embodiments of the present invention will be apparent to those skilled in the art. The present invention will now be exemplified by way of illustration with

reference to the following figures:

Figure 1 Positional Cloning of *Mlo*. The *Mlo* locus has been mapped with increasing precision on the long arm of barley chromosome 4 using morphological, RFLP and AFLP markers. The upper part of the figure presents the genetic linkage maps of these markers relative to *Mlo*. All genetic distances are indicated in centiMorgan (cM) based on multi-point linkage analysis except for genetic distances between AFLP markers which are calculated by two-point-estimates. The morphological marker map (Jørgensen, 1977) positions *Mlo* at a distance of more than 20 cM to hairy leaf sheath (*Hs*) and glossy sheath/spike (*gsl*). The RFLP marker map is based on the analysis of 257  $F_2$  individuals derived from the cross Carlsberg II *Mlo* Grannenlose Zweizeilige *mlo-11*. The previously published RFLP map (Hinze et al., 1991) of the same cross was based on only 44  $F_2$  individuals. The gene was delimited to a 2.7 cM interval bordered by markers bAO11 and bAL88. AFLP markers were identified and mapped as described in Experimental procedures. Their genetic distance to *Mlo* is based on the cross Ingrid *Mlo* x BC<sub>7</sub>Ingrid *mlo-3*. The crucial result of the AFLP analysis has been the identification of two markers, Bpm2 and Bpm9, defining an 0.64 cM interval containing the *Mlo* locus and one marker (Bpm16) cosegregating with *Mlo* on the basis of more than 4,000 meiotic events. Marker Bxm2 which is located 0.1 cM telomeric to *Mlo* was derived from BAC F15 template DNA (see below). One YAC clone, YAC YHV303-A6, containing the

5 cosegregating marker Bpm16 and two flanking loci (Bpm2 and Bpm9), is shown in the middle section of the figure. The position of marker Bpm9 was only roughly estimated within the YAC clone as indicated by the arrow. The insert of BAC F15 represents a 60 kb subfragment of this YAC as indicated in the lower part of the Figure. After the identification of AFLP marker Bpm2 in BAC F15, marker Bxm2 was discovered and positioned 0.1 cM in telomeric orientation of M10. The approximate physical position of AFLP markers Bpm2, Bpm16, and Bxm2 (spanning an interval of approximately 30 kb) as well as the location of some rare occurring restriction sites are indicated. Dashed lines below the schematic representation of BAC F15 DNA show the position of the largest established DNA sequence contigs. The structure of the M10 gene is given schematically in the bottom line of the Figure. Exons are highlighted by black boxes. Positions of mutational events are indicated for the eleven tested *m10* alleles. Mutant alleles carrying deletions in their nucleotide sequence are marked with a  $\Delta$ ; the remaining mutant alleles represent single nucleotide substitutions resulting in amino acid exchanges in each case.

Figure 2 shows an M10 coding sequence and encoded amino acid sequence according to the present invention. The amino acid sequence predicted from DNA sequences of RT-PCR products from Ingrid M10 are shown. Nucleotide numbers are given according to translational start site.

Figure 3 Northern Blot Analysis of M10 Transcript Accumulation. Total RNA (20  $\mu$ g) and poly(A)<sup>+</sup> RNA (5  $\mu$ g) of

0972237.142800



seven-day-old uninfected barley primary leaves of one wild type (cultivar Ingrid *Mlo*) and two mutant (BC Ingrid *mlo-1*, BC Ingrid *mlo-3*) cultivars were isolated, separated on a 1.2% formaldehyde gel and transferred to a nitrocellulose membrane (Hybond). The filter was probed under stringent conditions (Sambrook et al., 1989) with the radioactivity labelled full size RT-PCR product derived from Ingrid *Mlo* (Figure 7). A clear signal is detected only in the lanes containing poly(A)<sup>+</sup> RNA. The signal corresponds to a size of approximately 2 kb.

Figure 4 Southern Blot Analysis of Intragenic Recombinants derived from *mlo* heteroallelic crosses. The alleles of two RFLP markers flanking *Mlo* on opposite sides of either susceptible F<sub>2</sub> individuals or homozygous susceptible and homozygous resistant progeny were determined by Southern blot analysis. Plant DNA (10 µg) of the individuals were digested with *Pst* I (A) or *Hae* III (B) and hybridized with the radioactively labelled RFLP markers WG114 (upper panel; maps 3.1 cM in centromeric orientation to *Mlo*; see Figure 1) and ABG366 (lower panel; maps 0.7 cM in telomeric orientation to *Mlo*; see Figure 1) according to standard procedures (Sambrook et al., 1989).

A DNA of the parental lines *mlo-8* and *mlo-1* and two homozygous susceptible (S, *Mlo Mlo*) and two resistant (R, *mlo mlo*) progenies derived from two susceptible F<sub>2</sub> plants (designated 1 and 2) were tested. The DNAs in lanes S and R represent selection F<sub>3</sub> individuals from F<sub>3</sub> families obtained by selfing the susceptible F<sub>2</sub> individuals 1 and 2. Note that

susceptible  $F_2$  individuals are expected to be heterozygous at Mlo in this section scheme. Infection phenotypes were scored seven days after inoculation with the mlo avirulent isolate K1. DNA from a third susceptible individual of this heteroallelic cross (see Table 7) is not included in this Figure.

B DNA of the parental lines mlo-5 and mlo-1 and seven homozygous susceptible (S, Mlo Mlo) and seven resistant (R, mlo mlo) progeny derived from seven susceptible  $F_2$  plants (designated 1 to 7) were tested. The DNAs in lanes S and R represent selected  $F_3$  individuals from  $F_3$  families obtained by selfing the susceptible  $F_2$  individuals 1 to 7. DNA was analyzed from two further susceptible individuals of this heteroallelic cross only in the  $F_3$  generation (8\* and 9\*).

Figure 5 shows an alignment of genomic sequences covering the barley Mlo gene and a rice homologue isolated via crosshybridization with a barley gene specific probe. The top line shows the barley Mlo genomic DNA sequence (exon sequences underlined). The bottom line shows the rice genomic sequence containing the rice Mlo homologue.

Figure 6 shows an alignment of genomic sequences carrying the barley Mlo gene and a barley homologue isolated via crosshybridization with a barley gene specific probe. The top line shows the barley Mlo genomic DNA sequence (exon sequences underlined). The bottom line shows the genomic sequence containing the barley Mlo homologue.

Figure 7 Nucleotide and Deduced Amino Acid Sequence of the Barley Mlo cDNA. The nucleotide and the deduced amino acid

09/03/97 14:30

sequence are based on the combined data of RT-PCR and RACE obtained from experiments using RNA of cultivar Ingrid M10. The stop codon is marked by an asterisk, the putative polyadenylation signal is underlined and the detected termini of RACE products are indicated by arrows above the sequence. Positions of introns as indentified by comparison with corresponding genomic clones are labelled by triangles below the nucleic acid sequence. Six predicted transmembrane spanning helices according to the MEMSAT algorithm (Jones et al., 1994) are boxed in grey colour. A putative nuclear localization signal (K-K-K-V-R) and casein kinase II site (S-I-F-D) in the carboxy-terminal half of the protein are shown in bold type.

Figure 8 shows genomic sequence of rice (*Oryza sativa*) homologue including coding and flanking sequences.

Figure 9 shows genomic sequence of barley (*Hordeum vulgare*) homologue including coding and flanking sequences.

Figure 10 shows cDNA sequence of rice homologue.

Figure 11 shows cDNA sequence of barley homologue.

Figure 12 shows cDNA sequence of *Arabidopsis thaliana* homologue.

Figure 13 shows amino acid sequence of rice homologue.

Figure 14 shows amino acid sequence of barley homologue.

Figure 15 shows amino acid sequence of *Arabidopsis* homologue.

Figure 16 shows a pretty box of amino acid sequences of M10, barley, rice and *Arabidopsis* homologues.

All documents mentioned in this document are incorporated by reference.

EXAMPLE 1 - CLONING OF MLO OF BARLEY

5

*Targeted search for AFLP markers tightly linked to Mlo*

Efforts to increase the DNA marker density around Mlo were coordinated with attempts to construct a local high resolution genetic map. An alternative possibility would have been to extend the population size of the characterized cross Carlsberg II Mlo x Grannenlose Zweizeilige mlo-11 (Hinze et al., 1991) but it was felt to be advantageous to establish a high resolution map starting out from one of the available BC mlo lines and its recurrent parent line. Importantly, the donor parent of the BC line represents a different genetic background in comparison to the recurrent parent line. In this way, searching for linked AFLP markers could be started in parallel with generating a large mapping population from a cross between the same genetic lines. In addition, the BC line based cross allowed testing of colinearity of DNA markers in the vicinity of Mlo as determined from the cross Carlsberg II Mlo x Grannenlose Zweizeilige mlo-11 (Hinze et al., 1991). For the new cross a mlo-3 backcross (BC) line was used that had been backcrossed seven times into the genetic background Ingrid (BC, Ingrid mlo-3 ; Hinze et al., 1991). The line was previously characterized to carry a relatively small introgressed DNA segment on barley chromosome 4. In addition,

09722377.112600

the donor parent line Malteria Heda *mlo*-3 exhibits in comparison to DNA from the recurrent parent Ingrid polymorphisms with most of the identified RFLP loci linked to *Mlo*. Thus, by searching polymorphisms only between two DNA  
5 templates, from lines Ingrid *Mlo* and BC<sub>7</sub> Ingrid *mlo*-3, we hoped to increase the density of DNA markers with AFLPs around *Mlo* in a targeted manner.

The same two lines were crossed to establish a segregating population for high resolution mapping of DNA markers, formally  
10 representing an eighth backcross. F<sub>2</sub> individuals were scored for *mlo* resistance after powdery mildew inoculation with isolate K1 (virulent on Ingrid *Mlo* and avirulent on BC<sub>7</sub> Ingrid *mlo*-3 ). Initially, only a small fraction of the F<sub>2</sub> (77 individuals) was analyzed for recombination events with flanking RFLP markers.  
15 Analysis of four identified recombinants (designated 8-32-2, 7-38-4, 1-34-1, and 1-49-4) indicated colinearity of marker order in this cross compared to the previously analyzed cross Carlsberg II *Mlo* x Grannenlose Zweizeilige *mlo*-11 (Hinze et al., 1991). Several of the 77 F<sub>2</sub> seedlings which exhibited a  
20 susceptible phenotype and heterozygosity for the tested flanking DNA marker loci (bAO11, bAL88/2, and bAP91; Hinze et al., 1991) were grown to maturity to provide further selfed seed material segregating for *Mlo*/*mlo*-3 in the F<sub>3</sub> generation. In total, leaf material was harvested for high resolution  
25 marker mapping from 2,026 individuals derived from either the selfed F<sub>2</sub> or F<sub>3</sub> generation.

AFLP marker candidates were identified by testing all

0972237.13000

possible *Pst* I/*Mse* I primer combinations (1,024) extending into genomic sequences up to nucleotide positions +2 and -3, respectively. Similarly, almost 1,900 *Eco* RI/*Mse* I primer combinations (+3/+3) have been analyzed. Four DNA templates were included in this analysis: Ingrid Mlo, BC<sub>7</sub>, Ingrid mlo-3, a DNA pool of two phenotypically mlo resistant F<sub>2</sub> individuals, and a DNA pool of nine phenotypically susceptible F<sub>2</sub> individuals. The resistant and susceptible F<sub>2</sub> individuals which were included as DNA pools in the AFLP search had been selected from the above mentioned RFLP analysis of 77 F<sub>2</sub> segregants. The pooled F<sub>2</sub> DNA enabled us to control whether candidate polymorphisms detected between template DNA from the parents were heritable traits in the F<sub>2</sub>. All identified AFLP candidate markers have been re-examined with eight DNA templates: Ingrid Mlo, BC<sub>7</sub>, Ingrid mlo-3, DNA pools from individuals of three F<sub>2</sub> families which were phenotypically homozygous susceptible (MloMlo) according to K1 inoculation experiments; DNA of three resistant F<sub>2</sub> individuals. A total of 18 *Pst* I/*Mse* I and 20 *Eco* RI/*Mse* I primers were confirmed based on the selection procedure.

The number of identified AFLP markers made it useful to assign them first roughly to marker intervals based on the RFLP map around Mlo. It was hoped that this approach should enable both evaluation of the distribution of AFLPs among previously identified RFLP intervals close to Mlo and selection of a pair of flanking AFLP markers with which recombinants could be identified among the 2,026 segregants. For AFLP assignment we

used those four recombinants that had been identified with RFLP markers out of the above mentioned small sample of 77  $F_2$  segregants from Ingrid *Mlo* x BC<sub>7</sub> Ingrid *mlo*-3 (two recombinants in interval bAP91-bAL88, one in *Mlo*-bAO11, and one in bAO11-ABG366). A total of 18 AFLPs were found to be located within a genetic distance of approximately 3.5 cM including *Mlo*.

*Construction of a high resolution AFLP map around Mlo*

A two-step procedure was used to construct the high resolution AFLP map. First, all 2,026 segregants were screened for recombination events between two AFLP markers on opposite sides of *Mlo* and subsequently only the few identified recombinants were used to map all the identified AFLPs in the 3.5 cM target interval. AFLP markers Bpm1 and Bpm9 were chosen, detecting each allelic DNA fragments in Ingrid *Mlo* and BC<sub>7</sub> Ingrid *mlo*-3 and located on opposite sites of *Mlo* to screen DNA templates of the segregants for recombination events. Alternatively, the search for recombinants could have been carried out with the flanking RFLP markers bAO11 and bAL88. However, although the conversion into cleaved amplified polymorphic sites (CAPS) was successful for both markers, difficulties to display the alleles of both loci simultaneously from crudely purified genomic DNA were encountered. A total of 2,026 individuals ( $F_2$  or  $F_3$  segregants) were screened simultaneously with AFLP markers Bpm1 and Bpm9 and 98 recombinants were identified. AFLP analysis was subsequently carried out with each of the 98 DNA templates of the

recombinants to identify the alleles of each of the identified  
of AFLP loci. The recombinants have been selfed and  
inoculation experiments with powdery mildew isolate K1 were  
performed using at least 25 individuals of each recombinant  
5 family to deduce the alleles of the previous generation at the  
Mlo locus. The obtained data enabled the construction of a high  
resolution map around Mlo based on more than 4,000 meiotic  
events and a resolution of at least 0.025 cM derived via two-  
point estimates. The essential result has been the  
10 identification of a DNA marker cosegregating with Mlo (Bpm16)  
and two flanking markers (Bpm2 and Bpm9) at a distance of 0.25  
and 0.4 cM respectively (Figure 1).

*Construction of a large insert size barley YAC library,  
15 isolation of Bpm16 containing YACs, and physical delimitation  
of Mlo*

The genetic evidence indicates that mlo resistance is due  
to loss of function in the Mlo wild type allele. Therefore, it  
was decided to establish a large insert size YAC library from  
20 cultivar Ingrid Mlo into vector pYAC4 (Burke et al., 1987;  
Hieter, 1990). Megabase DNA suitable for YAC cloning  
experiments was prepared in mg amounts from mesophyll  
protoplasts of five-day-old seedlings according to a modified  
protocol described by Siedler and Graner (1991). The DNA was  
25 partially digested with Eco RI in the presence of Eco RI  
methyltransferase to obtain DNA fragments after preparative  
pulsed-field gel electrophoresis (PFGE) in the size range of



500-600 kb. After ligation with *Eco* RI digested pYAC4, the DNA was transformed into yeast strain AB1380 and colonies carrying recombinant pYAC4 DNA were selected on solidified synthetic complete medium lacking tryptophan and uracil (Sherman *et al.*, 1986). Forty randomly selected yeast colonies were tested for the presence of barley DNA using labelled barley genomic DNA in Southern experiments. The size of the YAC inserts was found after PFGE separations to vary between 500 and 800 kb. On average a genetic distance of 0.2 cM was expected to be represented on the individual recombinant YAC clone. A total of ~40,000 clones representing four barley genome equivalents have been generated.

Four YAC clones (designated 303A6, 322G2, 400H11, and 417D1) have been isolated with marker Bpm16 cosegregating with *Mlo*. Their insert size was determined by PFGE to be 650, 710, 650, and 820 kb respectively. AFLP analysis had shown that three of these clones (303A6, 322G2, and 417D1) contain also both flanking marker loci whereas clone 400H11 contains only loci Bpm16 and Bpm2. These findings strongly suggested that the *Mlo* gene had been physically delimited on recombinant YAC clones 303A6, 322G2, and 417D1.

YAC 303A6 was chosen for subcloning experiments into BAC vector pECSBAC4 containing a unique *Eco* RI site (Shizuya *et al.*, 1992; the vector pECSBAC4 is described by Frijters and Michelmore, 1996; submitted). Total yeast DNA of this clone was partially digested with *Eco* RI to obtain DNA fragments with an average size of 50 kb and ligated into *Eco* RI digested and

dephosphorylated BAC vector. Bacterial colonies containing YAC 303A6-derived DNA in pECSBAC4 were identified by replica colony hybridization experiments. One set of colony containing membranes was hybridized with labelled yeast AB1380 DNA and the replica set was hybridized with labelled PFGE-purified YAC303A6 DNA. Recombinant BAC clones containing the AFLP locus Bpm16 were subsequently identified using the cloned 108 bp Pst I/Mse I genomic Bpm16 fragment as a probe in colony hybridization experiments.

10 One BAC clone, BAC F15, containing an insert of ~ 60 kb was chosen for further detailed studies. It was found that the recombinant BAC clone contained in addition the AFLP marker locus Bpm2, but not Bpm9. At this point the BAC F15 insert DNA indicated successful physical delimitation in telomeric  
15 orientation but it was an open question whether the insert would contain bordering sequences in centromeric direction. Instead of constructing a BAC contig between Bpm 16 and Bpm9, the option to develop new polymorphic markers from BAC F15 was chosen. An allelic Xba I/Mse I polymorphism (designated Bxm2)  
20 was identified between the parental lines Ingrid M10 and BC, Ingrid m10-3.

An analysis of the 25 recombinant individuals carrying recombination events within the M10 containing interval Bpm2-Bpm9 enabled mapping of Bxm2 in centrometric orientation at a  
25 distance of 0.1 cM from M10. Only four out of the 16 available recombinants in the interval Bpm9-M10 and none of the 9 recombinants in the interval M10-Bpm2 were found to exhibit a

09722377.11500

recombination event between Bxm2 and *Mlo*. It was concluded that *Mlo* had been physically delimited on BAC F15 between marker loci Bpm2 and Bxm2 (Figure 1).

5 *Identification of the Mlo gene and mlo mutants*

A random sequencing project was initiated to determine sequence contigs of the ~60 kb insert of BAC F15 before marker Bxm2 was identified and shown to delimit the gene in telomeric orientation. In parallel, a physical map was generated  
10 (Figure 1). The physical map indicated that the flanking markers Bpm2 and Bxm2 are physically separated by ~30 kb. The sequence contigs were searched for regions of high coding probability using the UNIX versions of the STADEN program package. Only one sequence contig of almost 6 kb, including the  
15 cosegregating marker Bpm16, revealed an extensive region of high coding probability.

RT-PCR reactions were performed with total leaf RNA derived from cultivar Ingrid *Mlo* using a series of primers deduced from regions which indicated high coding probabilities  
20 and obtained in each case a distinct amplification product. Sequencing of the largest RT-PCR products revealed a single extensive open reading frame of 1,602 bp (Figure 2). The deduced putative protein of 533 amino acids has a molecular weight of 60.4 kDal. The ~1.7 kb RT-PCR product was used as a  
25 hybridization probe and detected a single RNA transcript of ~1.9 kb length. (Figure 3). A comparison of the genomic sequence and the largest RT-PCR fragment reveals 12 exons and

11 introns, each flanked by the characteristic splice site sequences (Figure 1).

Because marker Bpm16 is located at the 3' end of the above described gene (exon 11) and cosegregates with the *Mlo* locus, we started a direct PCR sequencing of the various available mutagen-induced *mlo* resistance alleles. We identified in 14 out of 15 tested mutant alleles nucleotide alterations which result either in single amino acid alterations, deletions or frame shifts of the wild type sequence (Table 1). We suspect that mutant allele *mlo*-2 is located within the promoter- or 5' untranslated sequences. The region is notoriously difficult to be sequenced via direct PCR sequencing from genomic DNA templates but experiments using a series of nested primers are likely to solve this problem. In summary, the comparative sequencing of genomic DNA from various mutant *mlo* lines and their respective *Mlo* wild type cultivars provided strong evidence that *Mlo* has been identified.

#### 20 *Intragenic recombinants*

It had been the intention to provide a chain of evidence for the molecular isolation of *Mlo* which did not rely upon complementation experiments via transgenic barley plants. We had chosen to develop an unusual genetic tool to confirm that the identified gene represented *Mlo*. It was reasoned that if the mutations observed in the above described gene caused resistance to the powdery mildew fungus, recombination events

between mutant allele sites should restore wild type sequences. It was predicted that those intragenic recombinants would exhibit susceptibility upon powdery mildew attack.

A crossing scheme was devised involving *mlo* resistance alleles *mlo-1*, *mlo-5*, and *mlo-8*. The mutant alleles originate from the genetic backgrounds Haisa (*mlo-1*) and Carlsberg II (*mlo-5* and *mlo-8*). Intermutant crosses were performed as shown in Table 2 generating in each case at least 10  $F_1$  plants.  $F_2$  populations were obtained by self-fertilization.  $F_2$  seedlings were screened for rare susceptible individuals after inoculation with powdery mildew isolate K1 which is virulent on each of the parental *Mlo* wild type cultivars. Susceptible  $F_2$  individuals were identified with a frequency of  $\sim 6 \times 10^{-4}$ . In contrast, if comparable numbers of progenies from selfings of each of the *mlo* mutants were tested for resistance to K1, no susceptible seedling was identified. This finding strongly indicated that the majority of the susceptible individuals derived from the intermutant crosses were not due to spontaneous reversion events of the mutant *mlo* alleles.

Inheritance of the susceptible  $F_2$  individuals was tested after selfing in  $F_3$  families. Each of the  $F_2$  individuals segregated susceptible and resistant  $F_3$  individuals indicating heterozygosity for alleles conferring resistance/susceptibility in the  $F_2$ . Homozygous susceptible  $F_3$  progeny were isolated for the majority of susceptible  $F_2$  individuals by selfing of  $F_3$  individuals and subsequent identification of  $F_4$  families in which only susceptible individuals were detected.

A molecular analysis of the susceptible individuals has been performed using RFLP markers known to be tightly linked (< 3 cM) on each side of the *Mlo* locus (Figure 4). RFLP marker WG114 maps in centromeric orientation relative to *Mlo*, marker ABG366 maps in the direction of the telomere. Detected RFLP alleles are shown for the intermutant crosses *mlo-8* x *mlo-1* (A) and *mlo-1* x *mlo-5* (B). DNA was analyzed either from susceptible  $F_2$  individuals (indicated by \*) or from homozygous susceptible (S) and homozygous resistant (R)  $F_3$  progeny obtained from selfed susceptible  $F_2$  individuals.

The homozygous susceptible  $F_3$  progeny from the susceptible  $F_2$  plant #1 of cross *mlo-8* x *mlo-1* (Figure 4) reveals the WG114 allele derived from the *mlo-1* parent in centromeric orientation next to *Mlo* and the ABG366 allele from the *mlo-8* parent in telomeric orientation to *Mlo*. The homozygous resistant  $F_3$  progeny from  $F_2$  plant #1 of this cross reveals in contrast only the flanking marker alleles derived from parent *mlo-1*. The finding strongly suggested that susceptibility in  $F_2$  plant #1 is caused by a cross-over type of recombination in the preceding meiosis of one chromosome which results in a restoration of the *Mlo* wild type allele whereas the second  $F_2$  chromosome of individual 1 contains a functionally unaltered *mlo-1* allele. The allelotypes of the RFLP loci of the homozygous susceptible  $F_3$  progeny from susceptible  $F_2$  plant #2 are identical to the one described above. However, flanking marker alleles from the homozygous resistant  $F_3$  progeny of this individual are in both cases

09/22/97 11:00

derived from the *mlo-8* parent. It is concluded that again a cross-over type of recombination restored one *Mlo* wild type allele in the susceptible  $F_2$  individual #2.

Nine susceptible  $F_2$  individuals were recovered from the cross *mlo-1* x *mlo-5* (Figure 4). For susceptible  $F_2$  individuals #1 to #7 both homozygous susceptible and homozygous resistant  $F_3$  progeny were analyzed at the DNA level. Note that only DNA from the heterozygous susceptible  $F_2$  individuals was analyzed in the case of individuals #8 and #9 (marked by a \*). The following allele patterns with respect to the flanking RFLP loci were observed: (i) homozygous resistant  $F_3$  progeny showed on both sides of *Mlo* either only the allelotypes of loci WG114 and ABG366 derived from the *mlo-1* parent (individuals #1, #3, #6, #7) or only the allelotypes derived from the *mlo-5* parent (individuals #2, #4, #5). (ii) Homozygous susceptible  $F_3$  progeny showed in contrast either only the allelotypes of both loci derived from the *mlo-5* parent (no. #3, #5, #6) or they showed different allelotypes on both sides of *Mlo* (individuals #1, #2, #4, #7). (iii) The homozygous susceptible  $F_3$  progeny with different allelotypes on both sides always contain in centromeric orientation the *mlo-1* derived WG114 allele and in telomeric orientation the *mlo-5* derived ABG366 allele. (iv) The heterozygous susceptible  $F_2$  individual #8 reveals on either side next to *Mlo* only the alleles derived from parent *mlo-5*. The heterozygous susceptible individual #9 reveals in centromeric orientation alleles derived from both parents *mlo-1* and *mlo-5* whereas only the *mlo-5* derived allele is detected in

telomeric orientation. A comprehensive interpretation of the data suggests that susceptibility in  $F_2$  individuals no. #1, #2, #4, #7, and #9 is caused by a cross-over type of recombination restoring the *Mlo* wild type allele. Non cross-over types of recombination may have restored the *Mlo* wild type allele in individuals no. #3, #5, #6, and #8.

A compilation of the detected flanking RFLP alleles of all isolated susceptible  $F_2$  individuals or homozygous  $F_3$  progeny is shown in Table 3. Note that individual #3 of the cross *mlo-8* x *mlo-1* is not shown in Figure 4. The compilation reveals that (i) cross-over types of recombination (CO) and non cross-over types of recombination (NCO) are found with a ratio of 7 : 5, (ii) cross-over types of recombination are resolved unidirectional, and (iii) NCO recombinants were not observed with parental *mlo-1*-linked RFLP alleles.

The CO type intragenic recombinants isolated from heteroallelic *mlo* crosses were used to test whether wild type sequences of the *Mlo* candidate gene had been restored. For the three relevant alleles *mlo-1*, *mlo-5*, *mlo-8* alleles candidate mutation sites have been identified (Table 1 and 4). Direct PCR sequencing of genomic DNA of susceptible intragenic recombinants derived from both heteroallelic crosses *mlo-1* x *mlo-8* and *mlo-1* x *mlo-5* revealed restoration of wild type sequences (Table 4). This observation strongly suggests that the intragenic cross over event occurred between nucleotide -1 and +483 in the former and +3 and +483 in the latter cross (according to translational start site). Thus, the molecular



analysis of seven intragenic recombinants from two heteroallelic crosses provides final proof that the above described candidate gene represents *Mlo*.

## 5 EXAMPLE 2 - HOMOLOGUES OF THE IDENTIFIED MLO GENE

The available expressed sequence tag (EST) databases of *Oryza sativa* (rice) and *Arabidopsis thaliana* were searched for homologous protein sequences. Five *Arabidopsis* cDNA clones were identified whose deduced amino acid sequences show substantial similarity to the *Mlo* protein. Remarkable is cDNA clone 205N12T7 which reveals a chance probability of  $1.2 \times 10^{-45}$ . In addition, at least one significant homologue was found in rice (OSR16381A).

A rice BAC library (Wang et al., 1995) has also been screened with a labelled barley genomic fragment containing *Mlo*. A BAC clone containing an insert of ~23 kb was isolated. Subsequent subcloning enabled isolation of a 2.5 kb *Pst* I genomic rice fragment showing strong cross-hybridization with the barley *Mlo* gene probe. DNA sequencing of this fragment revealed remarkable DNA sequence similarities within exon sequences of the barley *Mlo* gene (Figure 5).

Finally, a 13 kb  $\lambda$  genomic barley clone derived from cultivar Igri (Stratagene) was isolated with a labelled barley genomic fragment containing *Mlo*. The nucleotide sequence derived from a subcloned 2.6 kb *Sac* I fragment reveals again extensive sequence similarities to the *Mlo* gene (Fig. 6). The

location of the barley *Mlo* homologue within the genome is not within BAC F15 DNA.

In summary, there is conclusive evidence for *Mlo* homologues both in a monocotyledonous and a dicotyledonous plant species.

### Discussion

Any speculation as to mode of action of *Mlo* and *mlo* nucleic acid and polypeptides should provide no limitation on the nature or scope of any aspect or embodiment of the present invention.

In plants, resistance to pathogens is frequently determined by dominant resistance genes, whose products are assumed to recognize pathogen-derived avirulence gene products. This mode of pathogen defence follows Flor's gene-for-gene hypothesis (Flor, 1971). Recently, several 'gene-for-gene' type resistance genes have been molecularly isolated (Martin et al., 1993; Bent et al., 1994; Jones et al., 1994; Mindrinos et al., 1994; Whitham et al., 1994; Grant et al., 1995; Lawrence et al., 1995; Song et al., 1995). The surprising finding is that the deduced proteins share remarkable similar structural domains although they trigger resistance reactions to pathogens such as viruses, fungi, and bacteria (Dangl, 1995; Staskawicz et al., 1995). The isolated genes code for proteins that either contain a leucine-rich region (LRR), with or without an attached nucleotide binding site (NBS), indicative of ligand-binding and protein-protein interaction or encode a simple

serine/threonine kinase. A structural combination of LRR and the kinase domain has been reported in the deduced protein from the rice *Xa21* resistance gene (Song et al., 1995). The structural similarity of resistance genes in 'gene-for-gene' defence makes the existence of a common underlying resistance mechanisms likely.

Resistance mediated by recessive resistance alleles of the *Mlo* gene differs in various aspects from 'gene-for-gene' resistance (see introductory comments above). The molecular isolation of the *Mlo* gene and the sequencing of various mutation-induced *mlo* alleles described here, confirms previous interpretations from combined mutational and Mendelian genetic studies (Hentrich, 1979; Jørgensen, 1983). It is concluded that defective alleles of the *Mlo* locus mediate broad spectrum resistance to pathogens such as the powdery mildew pathogen. This is inconsistent with the involvement of a specific recognition event of a pathogen-derived product as has been proposed for race-specific resistance genes.

Pleiotropic effects of *mlo* alleles have provided some clues towards the development of a molecular concept of the observed broad spectrum resistance response.

Firstly, aseptically grown *mlo* plants exhibit at a high frequency a spontaneous formation of cell wall appositions (CWAs) in leaf epidermal cells (Wolter et al., 1993). Those CWAs are usually formed in response to attempted pathogen penetration directly beneath the fungal apressorium. CWAs are believed to form a physical barrier against pathogen ingress

and have been implicated repeatedly in *mlo* mediated resistance (Bayles, 1990).

Secondly, at a later stage, the plants develop macroscopically detectable leaf necrotic flecks. The spontaneous leaf necrosis response has been extensively studied with a unique collection of 95 chemically-induced *mlo* alleles (Hentrich, 1979). The alleles were classified as either showing a gradually different infection phenotype upon infection of a mixture of nine powdery mildew isolates. Those *mlo* alleles which give rise to an intermediate infection phenotype (i.e. development of a considerable number of sporulating fungal colonies upon inoculation) showed no detectable spontaneous leaf necrosis whereas the category of the most effective resistance alleles exhibits pronounced necrosis in the absence of the pathogen. Thus, there is solid evidence that the former category of *mlo* alleles retain residual wild type allele activity and those alleles appear to exhibit no detectable spontaneous leaf necrosis.

Thirdly, a constitutive expression of defence-related genes has been observed in *mlo* seedlings grown under mildew-free conditions - in primary leaves when 10-11 days old; this includes genes of the PR-1 family, chitinases and peroxidases.

We have shown that *mlo* in barley confers increased resistance to different types of yellow rust (*Puccinia struiformis*) when a one to one mixture of talcum powder and spores were aviblown onto leaves of *mlo* barley plants after onset of constitutive expression of defence related genes (10-

11 day old *mlo* seedlings).

Thus, it appears that multiple defence-associated responses are constitutively expressed in *mlo* plants.

The temporal relationship of these events is interesting:

5 the onset of constitutive defence-related transcript accumulation is detected in 11 day-old seedlings and precedes CWA formation which is followed by the appearance of macroscopically visible leaf necrosis. Importantly, however, *mlo* resistance can be experimentally tested as early as in five  
10 day-old seedlings and is fully functional at this time. We conclude that the Mlo protein has a negative regulatory function in plant defence and that plants with a defective protein are 'primed' for the onset of defence responses.

The deduced amino acid sequence of Mlo reveals no  
15 significant homologies to any of the described plant resistance genes so far, supporting the idea of a distinct molecular resistance mechanism. The Mlo gene shows also no striking similarities to any characterized plant or mammalian gene sequence in the various data bases. However, highly significant  
20 homologous sequences have been identified in the EST and genomic databases both from rice and *Arabidopsis thaliana* (Table 5 and Figure 5). This strongly suggests that the Mlo protein represents a member of a novel protein family. A putative nuclear localization motif (NLS) is found within exon  
25 12 providing indication of nuclear localization of the protein (KEKKKVR; Nigg et al., 1991). The significance of this motif is supported by a casein kinase II motif located 14 amino acids

into direction of the NH<sub>2</sub>-terminus (SIFD; Rihs et al., 1991). Functional tests may examine the putative subcellular localization of the Mlo protein.

Mutations have been described also in other plant species in which defence responses to pathogens appear to be constitutively expressed (Walbot et al., 1983; Pryor, 1987; Jones, 1994). It has been suggested that this class of mutants, termed lesion mimics (*Les*) or necrotic mutants (*nec*), affect the control of plant defence responses. Recessively inherited lesion mimic mutants have been systematically analysed in *Arabidopsis thaliana* (Greenberg and Ausubel, 1993; Dietrich et al., 1994; Greenberg et al., 1994; Weymann et al. 1995). The affected genes have been designated *acd* (accelerated cell death; *acd1* and *acd2*) or *lsd* (lesions simulating disease resistance response; *lsd1* to *lsd7*).

Each of the mutants exhibits, in the absence of pathogens, HR characteristics such as plant cell wall modifications and the accumulation of defence-related gene transcripts. Leaves of the *acd2* mutant have been shown to accumulate high levels of salicylic acid and of the *Arabidopsis* phytoalexin, camelexin (Tsuji et al., 1992). Importantly, *acd* and *lsd* mutants exhibit elevated resistance to a bacterial (*P. syringae*) and fungal (*P. parasitica*) pathogen. The *lsd1* mutant is exceptional in that it confers heightened pathogen resistance at a prelesion state, in contrast to the other defective loci which exhibit elevated pathogen resistance only in the lesion-positive state. In this respect, *lsd1* resembles the *mlo* mutants in barley. Another

striking feature of *lsd1* is the indeterminate spread of lesions in contrast to the other mutants where lesion growth is determinate.

## 5 EXPERIMENTAL PROCEDURES

### Plant Material

A compilation of the *mlo* mutants and their mother varieties analyzed in this study has been described by Jørgensen (1992) [*mlo*-1, *mlo*-3, *mlo*-4, *mlo*-5, *mlo*-7, *mlo*-8, *mlo*-9, *mlo*-10, *mlo*-11] and by Habekuss and Hentrich (1988) [mutants in cultivar Plena 2018 (*mlo*-13), 2034 (*mlo*-17), 2118]. Since mutant 2118 has not been assigned to an allele number so far, we designate the allele here as *mlo*-26, according to current numbering in the GrainGene database ([gopher://greengenes.cit.cornell.edu:70/77/.graingenes.ndx/index?mlo](http://gopher://greengenes.cit.cornell.edu:70/77/.graingenes.ndx/index?mlo)).

The high resolution map is based on a cross between Ingrid *Mlo* x BC<sub>7</sub> Ingrid *mlo*-3. F<sub>1</sub> plants were selfed generating a segregating F<sub>2</sub> population of approximately 600 plants. Phenotypically susceptible F<sub>2</sub> plants which showed heterozygosity for RFLP markers on opposite sites of *Mlo* were selfed and generated further segregants in the F<sub>3</sub> generation for high resolution mapping.

### Powdery Mildew Infection Tests

The fungal isolate K1 (Hinze et al., 1991) is virulent on

all cultivars used in this study carrying the *Mlo* allele and avirulent on all tested *mlo* genotypes. Plant growth and inoculation with *Erysiphe graminis* f sp *hordei* were carried out as described previously (Freialdenhoven et al., 1996). The genotype at *Mlo* of recombinants used for the high resolution map were determined after selfing and subsequent inoculation experiments in  $F_3$  or  $F_4$  families comprising at least 24 individuals.

#### AFLP Analysis

Genomic DNA for AFLP analysis was isolated according to Stewart and Via (1993). AFLP analysis was carried out with minor modifications as described by Vos et al. (1995). For screening of AFLP markers linked to *Mlo* we used the enzyme combinations Pst I/Mse I with amplification primers carrying +2 and +3 selective bases respectively in genomic sequences of amplified fragments. For Eco RI/Mse I amplification primers we used +3 and +3 selective bases respectively. A set of four DNA templates has been used: from the susceptible parent cultivar Ingrid *Mlo*, the resistant parent BC<sub>7</sub>Ingrid *mlo*-3, a pool of two resistant  $F_2$  individuals (*mlo*-3 *mlo*-3) and a pool of nine susceptible  $F_2$  individuals (*Mlo* *Mlo*) derived from the cross Ingrid *Mlo* x BC<sub>7</sub> Ingrid *mlo*-3. Amplified genomic fragments representing AFLP markers Bpm2, Bpm9, and Bpm16 (Figure 1) were cloned and sequenced as follows: gel pieces (fixed by vacuum drying to Whatman 3MM paper) containing the amplified genomic fragments were identified via autoradiography and subsequently



excised. 100  $\mu$ l water were added, boiled for 10 min. and after centrifugation 5  $\mu$ l of the supernatant were used as a template for non-radioactive reamplification (30 cycles) with the selective AFLP primers. Amplification products were isolated after agarose gel using a DNA isolation kit (Jetsorb, Genomed Inc., USA). DNA was reated with Klenow polymerase and T4 polynucleotide kinase and subsequently cloned in the EcoRV site of pBluescript SK (Stratagene). Sequencing reactions were performed using a dye terminator cycle sequencing reaction kit (Perkin Elmer) and resolved either on an ABI 373 or 377 (Applied Biosystems) automated sequencer.

*Barley YAC Library and BAC Sublibrary Construction of YAC YHV303-A6*

The YAC library of barley cultivar Ingrid was established using the pYAC4 vector (Burke et al., 1987; Kuhn and Ludwig 1994) and yeast strain AB 1380. Details of the library construction and its characterization will be described elsewhere. Screening for YAC clones containing marker Bpm16 was done by AFLP analysis. For construction of a BAC sublibrary of YAC YHV303-A6, total DNA of this yeast clone was used. After partial Eco RI digestion and preparative pulsed-field gel electrophoresis, DNA fragments in the size range of 50 kb were recovered and subcloned in the pECBAC4 vector. Clones carrying YHV303-A6 derived inserts were identified by a two-step colony hybridization procedure. First total labelled DNA of the non-recombinant yeast strain AB 1380 was used as a

probe to eliminate most of the clones carrying insert DNA derived from the host strain. In a subsequent hybridization step the remaining clones were probed with the labelled recombinant chromosome YHV303-A6 after enrichment by preparative pulsed-field gel electrophoresis.

#### *DNA Sequencing of BAC F15*

DNA of BAC F15 was isolated by an alkaline lysis large scale plasmid preparation according to Sambrook et al. (1989). 50 µg of purified DNA were nebulized by high pressure treatment with argon gas in a reaction chamber for 150 seconds. The ends of the sheared and reprecipitated DNA were blunt-ended by a T4 DNA polymerase-mediated fill in reaction. DNA fragments in the size range between 800 bp and 3 kb were isolated from agarose gels using a DNA isolation kit (Jetsorb, Genomed Inc., U.S.A.), subcloned into the pBluescript SK vector (Stratagene) and propagated in *E. coli* DN5α. Clones carrying BAC F15 derived inserts were selected by hybridization using the sheared DNA of BAC F15 as a probe. Sequencing reactions were performed as described above. Evaluation of the sequencing data, construction of sequence contigs, and estimation of coding propabilities were done by means of the STADEN software package for Unix users (4th edition, 1994). Assessment of coding probabilities was based on a combined evaluation of uneven positional base frequencies, positional base preference and barley codon usage in the investigated contigs. Homology searches were done using the BLAST software.

*PCR-based Sequencing of Alleles at Mlo*

Plant chromosomal DNA for this purpose was isolated according to Chunwongse et al. (1993). DNA sequences of *Mlo* alleles of the different barley varieties, *mlo* mutants, BC lines, and intragenic recombinants used in this study were obtained by PCR-based sequencing. Seven overlapping subfragments of the gene (each 400 bp-600 bp in length) were amplified by PCR (35 cycles, 60°C annealing temperature) using sets of specific primers. After preparative agarose gel electrophoresis and isolation of the amplification products using the Jetsorb kit (Genomed Inc., U.S.A.) fragments were reamplified to increase specificity. The resulting products were subsequently purified from nucleotides and oligonucleotides (Jetpure, Genomed Inc., U.S.A.) and used as a template in DNA sequencing reactions (see above). All DNA sequences of mutant alleles and corresponding regions of the parental lines and the intragenic recombinants were derived from both strands and confirmed two times in independent sets of experiments. In addition, mutant alleles *mlo*-1, *mlo*-3, *mlo*-4, *mlo*-5, *mlo*-7, *mlo*-8, *mlo*-9, and *mlo*-10 were also verified in the corresponding BC lines in cultivar Ingrid.

*RT-PCR and Rapid Amplification of cDNA Ends (RACE)*

RT-PCR was performed using the SUPERSSCRIPT preamplification system for first strand cDNA synthesis (Gibco BRL). Total RNA (1 µg) of seven-day-old primary barley leaves (cultivar Ingrid) served as template. First strand cDNA

synthesis was primed by an oligo(dT) primer. The putative coding region of the *Mlo* gene was subsequently amplified using oligonucleotides 25L (GTGCATCTGCGTGTGCGTA) and 38 (CAGAAACTTGTCTCATCCCTG) in a single amplification step (35 cycles, 60°C annealing temperature). The resulting product was analyzed by direct sequencing. 5'- and 3'-ends of the *Mlo* cDNA were determined by RACE (Frohman et al., 1988) using the MARATHON cDNA amplification kit (Clontech). Corresponding experimental procedures were mainly carried out according to the instructions of the manufacturer. To obtain specific RACE products, two consecutive rounds of amplification (35 cycles, 55°C annealing temperature) were necessary. For this purpose, two sets of nested primers were used in combination with the adapter primers of the kit: oligonucleotides 46 (AGGGTCAGGATCGCCAC) and 55 (TTGTGGAGGCCGTGTTCC) for the 5'-end and primers 33 (TGCAGCTATATGACCTTCCCCCTC) and 37 (GGACATGCTGATGGCTCAGA) for the 3'-end. RACE products were subcloned into pBluescript SK (Stratagene). Ten 5'-end and eight 3' end clones were chosen for DNA sequence analysis.

The term "AFLPs" is used herein to refer to "AFLP markers".

Table 1 summarizes the identified mutation sites of various mutants within the *Mlo* gene. The origin, the mutagen and the predicted effect of the mutation at the amino acid level are indicated.

5 Table 2 shows the results of heteroallelic *mlo* crosses and selfings of the respective *mlo* lines to isolate intragenic recombination events.

10 Table 3 summarizes the genotypes at flanking RFLP markers in susceptible  $F_2$  or homozygous  $F_3$  progeny from the intermutant crosses. CO and NCO indicate crossover type and non crossover type recombinants deduced from flanking molecular marker exchange. Table 3 summarizes DNA sequence analysis of susceptible intragenic crossover type recombinants (from homozygous susceptible  $F_3$  progeny) and the corresponding parental *mlo* mutant lines. Sequences flanking the identified mutation sites are shown.

15 Table 4 shows the results of direct PCR sequencing of genomic DNA of susceptible intragenic recombinants derived from both heteroallelic crosses *mlo-1* x *mlo-8* and *mlo-1* x *mlo-5*,  
20 revealing restoration of wild type sequences.

Table 5 shows several *Arabidopsis thaliana* and two rice expressed sequence tags (ESTs) with homology to the *Mlo* protein.

25 Table 5A show amino acid sequences, with "query" indicating part of the *Mlo* protein sequence to which homology has been found, with the predicted amino acid sequence of each identified EST marked with "subject".

Table 5B shows EST nucleotide sequences encoding the amino acid sequences shown in Table 5A. GenBank Accession number T22145 (definition 4153 Arabidopsis thaliana cDNA clone 97N8T7, NCBI Seq ID 932185), number T22146 (definition 4153 Arabidopsis thaliana cDNA clone 97N9T7, NCBI Seq ID 932186), number N37544 (definition 18771 Arabidopsis thaliana cDNA clone 205N12T7, NCBI Seq ID 1158686), number T88073 (definition 11769 Arabidopsis thaliana cDNA clone 155I23T7, NCBI Seq ID 935932) number H76041 (definition 17746 Arabidopsis thaliana cDNA clone 193P6T7, NCBI seq ID 1053292), number D24287 (rice cDNA partial sequence R1638\_1A, nID g428139) and D24131 (rice cDNA partial sequence R1408\_1A, nID g427985) are shown. The Arabidopsis sequences are from Newman et al. (1994) *Plant Physiol.* 106 1241-55. The rice sequences are from Minobe, Y. and Sasaki, T. submitted 2 Nov 1993 to DDBJ.

Allele	Mother Variety	Mutagen	Mutational Event at <i>Mlo</i>	Effect on Amino Acid Level
<i>mlo-1</i>	Haisa	X-rays	T <sup>484</sup> → A	trp <sup>182</sup> → arg
<i>mlo-3</i>	Malteria Heda	γ-rays	deletion of 2 nucleotides (1188-1189)	frame shift after phe <sup>395</sup>
<i>mlo-4</i>	Foma	X-rays	deletion of 11 nucleotides (478-488)	frame shift after trp <sup>159</sup>
<i>mlo-5</i>	Carlsberg II	EMS	G <sup>3</sup> → A	met <sup>1</sup> → ile <sup>a</sup>
<i>mlo-7</i>	Carlsberg II	EMS	G <sup>677</sup> → A	gly <sup>226</sup> → asp
<i>mlo-8</i>	Carlsberg II	EMS	A <sup>1</sup> → G	met <sup>1</sup> → val <sup>a</sup>
<i>mlo-9</i>	Diamant	EMS	C <sup>28</sup> → T	arg <sup>10</sup> → trp
<i>mlo-10</i>	Foma	γ-rays	deletion of 6 nucleotides (543-548)	2 amino acids (phe <sup>182</sup> , thr <sup>183</sup> ) missing
<i>mlo-12</i>	Elgina	NMU	C <sup>720</sup> → A	phe <sup>240</sup> → leu
<i>mlo-13</i>	Plena	EMS	T <sup>89</sup> → A	val <sup>30</sup> → glu
<i>mlo-16</i>	Alsa	EMS	G <sup>1917</sup> → A	alteration in 3' splice border of intron 9
<i>mlo-17</i>	Plena	EMS	C <sup>92</sup> → T	ser <sup>31</sup> → phe
<i>mlo-26</i>	Plena	EMS	T <sup>809</sup> → A	leu <sup>270</sup> → his
<i>mlo-28</i>	Nadja	NaN <sub>3</sub>	C <sup>665</sup> → T	...

Numbers of nucleotides and amino acids are given according to the translational start site of the *Mlo* cDNA sequence.

Nucleotide number according to the translational start site of the genomic *Mlo* DNA sequence.

EMS = ethylmethane sulfonate, NMU = nitrosomethylurea, NaN<sub>3</sub> = sodium azide.

a Next start codon is at nucleotide positions 79-81 and is in frame with the coding sequence.

Table 2

F<sub>2</sub> progeny from intermutant crosses and selfings

Testcrosses	resistant	susceptible	frequency of susceptible F <sub>2</sub> progeny
<i>mlo-8</i> x <i>mlo-1</i>	5,281	3	$5,7 \times 10^{-4}$
<i>mlo-5</i> x <i>mlo-1</i>	915	0	---
<i>mlo-5</i> x <i>mlo-1</i>	14,474	9	$6,2 \times 10^{-4}$
selfings	resistant	susceptible	
<i>mlo-1</i>	12,634	0	
<i>mlo-5</i>	5,498	0	
<i>mlo-8</i>	8,435	0	

097237 4400



Table 3. Genotypes at Flanking RFLP Markers in Susceptible Progeny Derived from Heteroallelic *mlo* Crosses

Testcross	Susceptible Plant	Parental Genotype in		Parental Genotype in Telomeric		Type of Recombination
		Centromeric Orientation to <i>Mlo</i> <sup>1</sup>	Orientation to <i>Mlo</i> <sup>2</sup>			
<i>mlo-8</i> x <i>mlo-1</i>	1	<i>mlo-1</i>	<i>mlo-8</i>		CO	
	2	<i>mlo-1</i>	<i>mlo-8</i>		CO	
	3	<i>mlo-8</i>	<i>mlo-8</i>		NCO	
<i>mlo-1</i> x <i>mlo-5</i>	1	<i>mlo-1</i>	<i>mlo-5</i>		CO	
	2	<i>mlo-1</i>	<i>mlo-5</i>		CO	
	3	<i>mlo-5</i>	<i>mlo-5</i>		NCO	
	4	<i>mlo-1</i>	<i>mlo-5</i>		CO	
	5	<i>mlo-5</i>	<i>mlo-5</i>		NCO	
	6	<i>mlo-5</i>	<i>mlo-5</i>		NCO	
	7	<i>mlo-1</i>	<i>mlo-5</i>		CO	
	8*	<i>mlo-5</i>	<i>mlo-5</i>		NCO	
	9*	<i>mlo-1</i> + <i>mlo-5</i>	<i>mlo-5</i>		CO	

<sup>1</sup>deduced from alleles of RFLP marker WG114 (see Figure 1)<sup>2</sup>deduced from alleles of RFLP marker ABG366 (see Figure 1)

CO = cross over type, NCO = non cross over type of recombination

\* Genotypes of flanking RFLP markers have been determined in heterozygous susceptible F<sub>2</sub> individuals; in all other cases homozygous susceptible F<sub>3</sub> progeny derived from the susceptible F<sub>2</sub> individuals were tested

Table 4. Restoration of *Mlo* Wild Type Sequences by Intragenic Recombination Events

		Nucleotide Sequences Flanking Mutant Sites <sup>1</sup>	
Genotypes		Nucleotides -3 to +3	Nucleotides 481 to 486
Haisa	<i>Mlo</i>	CCGATG	AATGGG
	<i>mlo-1</i>	CCGATG	AAAGGG
Carlsberg II	<i>Mlo</i>	CCGATG	AATGGG
	<i>mlo-5</i>	CCGATA	AATGGG
	<i>mlo-8</i>	CCGGTG	AATGGG
	Intragenic recombinant		
	<i>mlo-1</i> x <i>mlo-8</i>	1 CCGATG	1 AATGGG
		2 CCGATG	2 AATGGG
	<i>mlo-1</i> x <i>mlo-5</i>	1 CCGATG	1 AATGGG
		2 CCGATG	2 AATGGG
		4 CCGATG	4 AATGGG
		7 CCGATG	7 AATGGG
		9 CCGATG	9 AATGGG

<sup>1</sup> Numbers of nucleotides are given according to the translational start site (see Figure 2).

TABLE 5A

>EM EST1:AT1452 T22145 4153 Arabidopsis thaliana cDNA clone 97N8T7. 11/95  
Length = 382

## Plus Strand HSPs:

Score = 248 (115.9 bits), Expect =  $2.9e-27$ , P =  $2.9e-27$   
Identities = 47/100 (47%), Positives = 67/100 (67%), Frame = +2

Query: 242 KYIKRSMEDDFKVVVGISLPLWGVAILTLFLDINGVGTLIWISFIPLVILLCVGTKLEMI 301  
KY+ R++EDDFK VVGIS LW ++ L++NG T WI+FIP +LL VGTKLE +  
Sbjct: 2 KYMMRALEDDFKQVVGISWYLWXFVVFIXLLNVNGWHTYFWIAFIPFXLLAVGKLEHV 181

Query: 302 IMEMALEIQDRASVIKGA PVVEPSNKFFWFH RPDWVLF 341  
I ++A E+ ++ I+G VV+P . + FWF +P VL+ I  
Sbjct: 182 IAQLAHEVAEKHVAIEGDLVVKPXXEHFWFSKPQIVLYLI 301

>EM EST1:AT1462 T22146 4154 Arabidopsis thaliana cDNA clone 97N9T7. 11/95  
Length = 390

## Plus Strand HSPs:

Score = 212 (99.1 bits), Expect =  $4.2e-26$ , Sum P(2) =  $4.2e-26$   
Identities = 41/83 (49%), Positives = 58/83 (69%), Frame = +2

Query: 242 KYIKRSMEDDFKVVVGISLPLWGVAILTLFLDINGVGTLIWISFIPLVILLCVGTKLEMI 301  
KY+ R++EDDFK VVGIS LW ++ L L++NG T WI+FIP +LL VGTKLE +  
Sbjct: 2 KYMMRALEDDFKQVVGISWYLWXFVVFILLNVNGWHTYFWIAFIPFALLAVGKLEHV 181

Query: 302 IMEMALEIQDRASVIKGA PVVEP 324  
I ++A E+ ++ I+G VV+P  
Sbjct: 182 IAQLAHEVAEKHVAIEGDLVVKP 250

Score = 52 (24.3 bits), Expect = 1.9, Sum P(2) = 0.85  
Identities = 9/32 (28%), Positives = 16/32 (50%), Frame = +2

Query: 18 WAVAVVFAAMVLVSVLMEHGLHKLGHWFQHRH 49  
W + FA ++ V +EH + +L H +H  
Sbjct: 122 WIAFIPFALLAVGKLEHVIAQLAHEVAEKH 217

Score = 49 (22.9 bits), Expect =  $4.2e-26$ , Sum P(2) =  $4.2e-26$   
Identities = 8/17 (47%), Positives = 12/17 (70%), Frame = +1

Query: 323 EPSNKFFWFH RPDWVLF 339  
E S++ EWF +P VL+  
Sbjct: 244 ETSDEHFWFSKPQXVLY 294

TABLE 5A cont'd

>EM EST1:AT54418 N37544 18771 Arabidopsis thaliana cDNA clone 205N12T7. 1/96  
Length = 585

Plus Strand HSPs:

Score = 277 (129.5 bits), Expect = 1.2e-45, Sum P(2) = 1.2e-45  
Identities = 51/96 (53%), Positives = 71/96 (73%), Frame = +1

Query: 236 SKFDFHKYIKRSMEDDFKVVVGISLPLWGVAILTLFLDINGVGTLIWISFIPLVILLCVG 295  
S+EDE KYI+RS+E DFK VV IS +W VA+L L + G+ + +W+ FIPLV++L VG  
Sbjct: 127 SRFDFRKYIQRSLEKDFKTVVEISPVIWFVAVLFLLTNSYGLRSYLWLPFIPLVVILIVG 306

Query: 296 TKLEMIIMEMALEIQDRASVIKAPVVEPSNKKFFWF 331  
TKLE+II ++ L IQ+ V++GAPVV+P + FWF  
Sbjct: 307 TKLEVIITKLGLRIQEEDVVRGAPVVQPGDDXFWF 414

Score = 121 (56.6 bits), Expect = 1.2e-45, Sum P(2) = 1.2e-45  
Identities = 25/45 (55%), Positives = 29/45 (64%), Frame = +1

Query: 196 SSTPGIRWVVAFFRQFFRSVTKVDYLTLRAGFINAHLSONSKFDF 240.  
S T W+V FFRQFF SVTKVDYL L GFI AH + ++ F  
Sbjct: 1 SKTRVTLWIVCFFRQFFGSVTKVDYLALXHGFI MAHFAPGNESRF 135

>EM EST1:AT04117 H76041 17746 Arabidopsis thaliana cDNA clone 193P6T7. 11/95  
Length = 476

Plus Strand HSPs:

Score = 210 (98.2 bits), Expect = 9.0e-36, Sum P(2) = 9.0e-36  
Identities = 43/86 (50%), Positives = 58/86 (67%), Frame = +1

Query: 196 SSTPGIRWVVAFFRQFFRSVTKVDYLTLRAGFINAHLSONSKFDFHKYIKRSMEDDFKVV 255  
++TP V FFRQFF SV + DYLT LR GF +AHL+ KF+F +YIK S+EDDFK+V  
Sbjct: 124 TTTPEXFNVGCFRQFFVSVERTDYLT LRHGFXSAHLAPGRKFNFQRYIKXSLEDDFKLV 303

Query: 256 VGISLPLWGVAILTLFLDINGVGTLI 281  
VGI LW ++ L + +GT++  
Sbjct: 304 VGIXPVLWASFVIFLAVQX\*WLG TIV 381

Score = 119 (55.6 bits), Expect = 9.0e-36, Sum P(2) = 9.0e-36  
Identities = 24/57 (42%), Positives = 32/57 (56%), Frame = +1

Query: 156 MRTWKKWETETTSLEYQFANDPARFRFTHQTSFVKRHLGLSSTPGIRWVVAFFRQFF 212  
+R WKKWE T S +Y F D +R R TH+TSFV+ H +T + V F + F  
Sbjct: 1 IRGWKKWEQXTLSNDYXFXIDHSRLRLTHETSFVREHTSFWTTPFXFNVGCFRQF 171

Score = 40 (18.7 bits), Expect = 1.2e-08, Sum P(2) = 1.2e-08  
Identities = 8/19 (42%), Positives = 10/19 (52%), Frame = +2  
Query: 269 TLFLDINGVGTLIWISFIP 287  
+L + NG G L W S P  
Sbjct: 344 SLLFNKXNGWGPLFWASVPP 400

TABLE 5A cont'd

>EM EST1:AT0739 T88073 11769 Arabidopsis thaliana cDNA clone 155I23T7. 11/95  
Length = 460

Plus Strand HSPs:

Score = 175 (81.8 bits), Expect = 1.2e-24, Sum P(2) = 1.2e-24  
Identities = 31/67 (46%), Positives = 43/67 (64%), Frame = +1

Query: 146 VITIALSRLKMRWKKWETETTSLEYQFANDPARFRFTHQTSFVKRHLGLSSTPGIRWVV 205  
++T A ++KMRTWK WE ET ++EYQ++NDP RFRF TSF +RHL S + +  
Sbjct: 4 IVTYAFGKIKMRTWKSWEETKTIEYQYSNDPERFRFARDTSFGRRLNFWSKTRVTLWI 183

Score = 121 (56.6 bits), Expect = 1.4e-14, Sum P(2) = 1.4e-14  
Identities = 25/45 (55%), Positives = 29/45 (64%), Frame = +1

Query: 196 SSTPGIRWVVAFFRQFFRSVTKVDYLTLRAGFINAHLSONSKFDF 240  
S T W+V FFRQFF SVTKVDYL L GFI AH + ++ F  
Sbjct: 157 SKTRVTLWIVCFRRQFFGSVTKVDYLALXHGFI MAHFAPGNESRF 291

Score = 75 (35.1 bits), Expect = 1.2e-24, Sum P(2) = 1.2e-24  
Identities = 14/21 (66%), Positives = 17/21 (80%), Frame = +1

Query: 236 SKFDFHKKYIKRSMEDDFKVVV 256  
S+FDK KYI+RS+ DFK VV  
Sbjct: 283 SRFDFRKYIQRSLXXDFKTVV 345

>EM EST5:OSR16381A D24287 Rice cDNA, partial sequence (R1638\_1A). 5/95  
Length = 400  
Plus Strand HSPs:

Score = 147 (68.7 bits), Expect = 1.9e-16, Sum P(2) = 1.9e-16  
Identities = 26/53 (49%), Positives = 35/53 (66%), Frame = +1

Query: 236 SKFDFHKKYIKRSMEDDFKVVVVGISLPLWGVAILTFLDINGVGTLIWISFIPL 288  
++F+F KYIKR +EDDFK VVGIS P W A+ + ++G L W S PL  
Sbjct: 202 TRFNFRKYIKRXLEDDEKTVVGISAPXWASALAIMLNFVHGHNLFWFSTXPL 360

Score = 45 (21.0 bits), Expect = 1.9e-16, Sum P(2) = 1.9e-16  
Identities = 9/15 (60%), Positives = 11/15 (73%), Frame = +2

Query: 287 PLVILLCVGKLEMI 301  
PL + L VGTKL+ I  
Sbjct: 356 PLXVTLAVGKTLQAI 400

>EM EST5:OSS1692A D39989 Rice cDNA, partial sequence (S1692\_1A). 11/94  
Length = 343

Plus Strand HSPs:

Score = 95 (44.4 bits), Expect = 0.00059, P = 0.00059  
Identities = 24/58 (41%), Positives = 31/58 (53%), Frame = +3

Query: 43 HWFQHRHKKALWEALEKMKAEMLVGFISLLLVTDQPIIAKICISEDAAADVWNECKR 100  
H + H+ L +A+EKMK E+ML+GFISLLL T I S+ PC R  
Sbjct: 3 HXSEKTHRNPLHKAMEKMKKEEMLLGFISLLLAATSRIISGICIDSKYNSNFSPECTR 176

TABLE 5B

100

## GenBank Accession Number T22145

```

1  caagtatatg atgcgcgctc tagaggatga tttcaaacaa gttgttggtg ttagttggta
61 tctttggntc tttgtcgta tcttttttct gctaaatgtt aacggatggc acacatattt
121 ctggatagca tttattccct ttnccttgct tcttgctgtg ggaacaaagt tggagcatgt
181 nattgcacag ttagctcatg aagttgcaga gaaacatgta gccattgaag gagacttagt
241 ggtgaaaccc ncanatgagc atttctgggt cagcaaacct caaattgttc tctacttgat
301 cccattttat cctctttccc agaatgcntt ttnagantgc nttttttntt tttggntttt
361 ggggtaanan annnggtttcg nc

```

## GenBank Accession Number T22146

```

1  caagtatatg atgcgcgctc tagaggatga tttcaaacaa gttgttggtg ttagttggta
61 tctttggntc tttgtcgta tcttttttct gctaaatgtt aacggatggc acacatattt
121 ctggatagca tttattccct ttgcttgct tcttgctgtg ggaacaaagt tggagcatgt
181 nattgcacag ttagctcatg aagttgcaga gaaacatgta gccattgaag gagacttagt
241 ggtgaaacct cagatgagca tttctgggtc agcaaacctc aaantgttct ctactngatc
301 cnctttatcc cccttccaga atgccttttt nangattcnn ntttttcctt nttgganntt
361 ttgggnnnnc aaacgggntt nggacctccg

```

## GenBank Accession Number N37544

```

1  agcaagacga gagtcacact atggattgtt tgttttttta gacagttctt tggatctgtc
61 accaaagttg attacttagc actaagnat ggtttcatca tggcgcatth tgctcccggg
121 aacgaatcaa gattcgattt ccgcaagtat attcagagat cattagagaa agacttcaaa
181 accgttggtg aaatcagtcg gggttatctg tttgtcgctg tgctattcct cttgaccaat
241 tcatatggat tacgttctta cctctgggtt ccattcattc cactagtcgt aattctaata
301 gttggaacaa agcttgaagt cataataaca aaattgggtc taaggatcca agaggaaggt
361 gatgtggtga gaggcgcccc agtgggtcag cctgggtgat accncttctg gtttngnaan
421 cacgnttcaa tnttttccnt antcacttng gcctttttan ggggtgaattt caacttcatn
481 ctttncctgg ggcgggatga ttcaatccaa naatnttccc ctgaagnctn caagtttggtg
541 cataggcttt nggtgggntt ttcaganttt nagtttggtc tnccc

```

TABLE 5B (Continued)

## GenBank Accession Number T88073

1 tgcattgtta cttatgcttt cggaaagatc aagatgagga cgtggaagtc gtgggaggaa  
 61 gagacaaaga caatagagta tcagtattcc aacgatcctg agaggttcag gtttgcnagg  
 121 gacacatctt ttgggagaag acatctcaat ttctggagca agacgagagt cacactatgg  
 181 attgtttgtt tttttagaca gttcttttga tctgtcacca aagttgatta cttagcacta  
 241 agncatgggt tcatcatggc gcatttttgc cccggtaacg aatcaagatt cgatttccgc  
 301 aagtatatcc agagatcatt agnnaagac ttcaaaaccg ttgtttgaaa tcagtcgggt  
 361 tatctgggtt gtcggctgtg ctattccnct tgaccaattc atatggntnc ggtnttncnc  
 421 tggtaaccatt attcncntag ggaatntaaa agttggcnga

## GenBank Accession Number H76041

1 attcgtggat ggaaaaagtg ggagcaagan acattatcta atgactatna gtttntctatt  
 61 gatcattcaa gacttaggct cactcatgag acttcttttg tnagagaaca tacaagtttc  
 121 tggacaacaa cncctttctn ctttaacgtc ggatgcttct ttaggcagtt ctttgtatct  
 181 gtngaaagaa ccgactactt gactctgcgc catggattca nctctgccca tttagctcca  
 241 ggaagaaagt tcaacttcca gagatatatc aaangatctc tcgaggatga tttcaagttg  
 301 gtagttggaa taagnccagt tctttgggca tcatttgtaa tcttccttgc tgttcaatgn  
 361 taatggctgg ggaccattgt tttgggcntc ggtaccgcct ntactcanaa ncccaggctt  
 421 ttggccaagg ttcaaggaat ttngggacaa tggggtagaa tcgtgggcnc atnngg

## GenBank Accession Number D24287

1 tcntnttttn ttttcgnntn cntccacccc tnnntnctc nancncttn nnnttatctc  
 61 tntntntntc ncntntcccn ncaccacccn ncgacgggcn tggactnngc cennngttcg  
 121 aggctgccca ctgncgtctg agacctacct tgn catttga cggcacngga cttcanttgc  
 181 tgctcacttt atctctacgg gactaggttc aattttcggg aatacatcaa aaggncactg  
 241 gaggacgatt ttaagacagt tgttggcatt agtgcacccn tatgggcttc tgcgttggcc  
 301 attatgctct tcaatgttca tggatggcat aacttggtct ggttctctac aatnccccct  
 361 gntagtaact ttagcagttg gaacaaagct gcaggctata

## GenBank Accession Number D24131

1 cagactacct gactttgagg cacggattca ttgctgctca tttatctcta gggactaggt  
 61 tcaattttcg gaaatacatc aaaaggtcac tggaggacga ttttaagaca gttgttggca  
 121 ttagtgcacc cttatgggct tctgcgttgg ccattatgct cttnaatgtt catggatggc  
 181 ataacttggt ctgggttctct acaatccccc ttgtagtaac tttagcagtt ggaacaaagc  
 241 tgcaggctat aattgcaatg atggctgttg aaattaaaga gaggcataca gtaattcaag  
 301 gaatgccggg ggtgaactca gtgat

## REFERENCES

1. Bayles, et al. (1990). *Physiol. Mol. Plant Pathol.* 36, 63-72.
2. Becker, et al. (1995). *Mol. and Gen. Genet.* 249, 65-73.
- 5 3. Bennet, et al. (1991). *Phil. Trans. R. Soc. London (Biol)* 334, 309-345.
4. Bent, et al. (1994). *Science* 265, 1856-1860.
5. Boyd, et al. (1995). *Plant. J.*, 7, 959-968.
6. Bourque, J. E. (1995). *Plant Science* 105, 125-149.
- 10 7. Burke, et al. (1987). *Science* 236, 806-812.
8. Cao, et al. (1992). *Plant Cell Rep.* 11, 585-591.
9. Christou, et al. (1991). *Bio/Technology* 9, 957-962.
10. Chunwongse, et al. (1993). *Ther. Appl. Genet.*, 86, 694-698.
- 15 11. Civardi (1994).
12. Cornejo, et al. (1993). *Plant Molecular Biology* 23, 567-581.
13. Dangl, J. L. (1995). *Cell* 80, 363-366.
14. Dangl, J. L., et al. (1996). *Plant Cell (in press)*.
- 20 15. Datta, et al. (1990). *Bio/Technology* 8, 736-740.
16. de Feyter, et al. (1996). *Molecular and General Genetics* 250, 329-338.
17. D'Halluin, et al. (1992). *Plant Cell* 4, 1495-1505.
18. Dietrich, et al. (1994). *Cell* 77, 565-577.
- 25 19. Doke, N. (1983). *Physiol. Plant Pathol.*, 23, 345-357.
20. Doke, N. and Ohashi, Y. (1988). *Physiol. Mol. Plant Pathol.*, 32, 163-175.



21. Dooner, H.K. and Kermicle, J.L. (1986). *Genetics*, 113, 135-143.
22. Dujon, B. (1996). *Trends Genet.*, 12, 263-270.
23. Flavell, R. B. (1994). *Proc. Natl. Acad. Sci. USA* 91, 3490-3496.
24. Flor, H. H. (1971). *Annu. Rev. Phytopathol.* 9, 275-296.
25. Freeling, M. (1978). *Genetics*, 89, 211-224.
26. Freialdenhoven, A., et al. (1996). *Plant Cell*, 8, 5-14.
27. Frohman, N.A., et al. (1988). *Proc. Natl. Acad. Sci.*, 85, 8998-9002.
28. Fromm, et al. (1990). *Bio/Technology* 8, 833-839.
29. Fugimoto, et al. (1993). *Bio/Technology* 11, 1151-1155.
30. Giovannoni, J.J., et al. (1991). *Nucl. Acids Res.*, 19, 6553-6558.
31. Gordon-Kamm, et al. (1990). *Plant Cell* 2, 603-618.
32. Görg, R., et al. (1993). *Plant J.*, 3, 857-866.
33. Grant, et al. (1995). *Science* 269, 843-846.
34. Greenberg, et al. (1993). *The Plant Journal* 4, 327-341.
35. Greenberg, et al. (1994). *Cell* 77, 551-563.
36. Habekuss, A. and Hentrich, W. (1988). *Tag. Ber., Akad. Lanwirtsch.-Wiss. DDR, Berlin*, 272: 229-237.
37. Hammond-Kosack, K.E. and Jones, J.D.G. (1996). *The Plant Cell (in press)*.
38. Hartmann, E., et al. (1989). *Proc. Natl. Acad. Sci.*, 86: 5786-5790.
39. Haselhoff, et al. (1988). *Nature* 334, 585-491.
40. Heath, M.C. (1980). *Annu. Rev. Phytopathol.*, 18: 211-236.

41. Hentrich, W. (1979). Arch. Züchtungsvorsch., Berlin 9, S. 283-291.
42. Heslop-Harrison, J.S. (1991). J. Cell Sci., 100, 15-21.
43. Hiei, et al. (1994). The Plant Journal 6, 271-282.
- 5 44. Hieter, et al. (1990). Cold Spring Harbor, NY, Col Spring Harbor Press.
45. Hinze, et al. (1991). Proc. Natl. Acad. Sci. USA 88, 3691-3695.
46. Jabs, T., et al. (1996). Science, 273, 1853-55.
- 10 47. Jones, D.T., et al. (1994). Biochemistry, 33, 3038-3049.
48. Jones, et al. (1994). Science 266, 789-793.
49. Jones, J. D. G. (1994). Current Biology 4, 749-751.
50. Jørgensen, J. H. (1977). Induced mutations against plant diseases (Crop Symposium Vienna) 533-547.
- 15 51. Jørgensen, J. H. (1983). Induced mutations for disease resistance in crop plants II (International Atomic Energy Agency, Vienna) 73-87.
52. Jørgensen, J. H. (1992). Euphytica 63, 141-152.
53. Jørgensen, J. H. (1994). Critical Reviews in Plant Sciences 13, 97-119.
- 20 54. Jørgensen, J.H. and Mortensen, K. (1977). Phytopathology, 67, 678-685.
55. Klein, P., et al. (1985). Biochim. Biophys. Acta, 815, 468-476.
- 25 56. Koga, H., et al. (1990). Can. J. Bot., 68, 2344-2352.
57. Koornneef, M., et al. (1983). Genet. Res. Camb., 41, 57-68.

58. Kosslak, R., et al. (1996). J. Hered (in press).
59. Kuhn, R.M. and Ludwig. R.A. (1994). Gene, 141, 125-127.
60. Künzel, G. (1982). Theor. Appl. Genet., 64, 25-29.
- 5 61. Koziel, et al. (1993). Biotechnology 11, 194-200.
62. Lamb, C.J. (1994). Cell, 76: 419-422.
63. Lawrence, et al. (1995). The Plant Cell 7, 1195-1206.
64. Levine, A., et al. (1994). Cell, 79: 583-593.
65. Li, et al. (1993). Plant Cell Rep. 12, 250-255.
- 10 66. Linde-Laursen, I., et al. (1982). Z. Pflanzenzüchtg., 81, 191-219.
67. Lundqvist, U., et al. (1991). Hereditas, 115, 227-239.
68. Martin, et al. (1993). Science 262, 1432-1436.
69. McClintock, B. (1984). Science, 226, 792-801.
- 15 70. Mindrinos, et al. (1994). Cell 78, 1089-1099.
71. Moore, et al. (1995). Current Biology 5, 737-739.
72. Mourad, G., et al. (1994). Mol. Gen. Genet., 243, 178-184.
73. Negassa, M. (1985). Hereditas, 102, 113-121.
- 20 74. Nigg, et al. (1991). Cell 66, 15-22.
75. Osborne, et al. (1995). Current Opinion in Cell Biology 7, 406-413.
76. Olsen, O., et al. (1993). Proc. Natl. Acad. Sci. USA, 90, 8043-8047.
- 25 77. Peng, et al. (1992). Theor Appl Genet 83, 855-863.
78. Peng, et al. (1991). International Rice Research Institute, Manila, Philippines. 563-574.

79. Pryor, A. J. (1987). Trends in Genetics 3, 157-161.
80. Raff, M.C. (1992). Nature, 356, 397-400.
81. Raff, M.C., et al. (1993). Science, 262, 695-700.
82. Rathore, et al. (1993). Plant Molecular Biology 21, 871-  
5 884.
83. Rihs, et al. (1991). EMBO J. 10, 633-639.
84. Ryerson, D.E. and Heath, M.C. (1996). Plant Cell, 8, 393-  
402.
85. Salamini, F., and Lorenzoni, C. (1970). Mol. Gen. Genet.,  
10 108, 225-232.
86. Sambrook, J., et al. (1989). Cold Spring Harbor  
Laboratory Press, New York.
87. Schmidt, R., et al. (1995). Science, 270, 480-483.
88. Sherman, et al. (1986). Cold Spring Harbor, NY, Cold  
15 Spring Harbor Laboratory Press.
89. Shimamoto, K. (1994). Current Opinion in Biotechnology 5,  
158-162.
90. Shimamoto, et al. (1989). Nature 338, 274-276.
91. Shirley, B.W., et al. (1992). Plant Cell, 4, 333-347.
- 20 92. Shizuya, et al. (1992). Proc. Natl. Acad. Sci. USA 89,  
8794-8797.
93. Siedler, et al. (1991). Molecular and General Genetics  
226, 117-181.
94. Søgaard, et al. (1988). Barley Genet. Newslett. 17, 120-  
25 134.
95. Somers, et al. (1992). Bio/Technology 10, 1589-1594.
96. Song, et al. (1995). Science 270, 1804-1806.

97. Stakman, E.C. (1915). *J. Ag. Research*, 4: 193-199.
98. Staskawicz, et al. (1995). *Science* 268, 661-667.
99. Stewart, C.N. and Via, L.E. (1993). *BioTechniques*, 14: 748-750.
- 5 100. Szostak, J.W., et al. (1983). *Cell*, 33: 25-35.
101. Tanksley, et al. (1995). *Trends in Genetics* 11, 63-68.
102. Thomas, C.M., et al. (1995). *Plant J.*, 8: 785-794.
103. Toriyama, et al. (1988). *Bio/Technology* 6, 1072-1074.
104. Tsuji, et al. (1992). *Plant Physiology* 98, 1304-1309.
105. Vasil, I. K. (1994). *Plant Molecular Biology* 25, 925-937.
106. Vasil, et al. (1992). *Bio/Technology* 10, 667-674.
107. Vos, et al. (1995). *Nucleic Acids Research* 23, 4407-4414.
108. Walbot, et al. (1983). New York, Plenum Press 431-442.
109. Walters, et al. (1992). *Plant Molecular Biology* 18, 189-200.
110. Wang, et al. (1995). *Plant Journal* 7, 525-533.
111. Weeks, et al. (1993). *Plant Physiology* 102, 1077-1084.
112. Weymann, et al. (1995). *The Plant Cell* 7, 2013-2022.
113. White, E. (1996). *Genes & Development*, 10: 1-15.
- 20 114. Whitham, et al. (1994). *Cell* 78, 1011-1115.
115. Wiberg, A. (1974). *Hereditas* 77, 89-148.
116. Wolter, et al. (1993). *Mol. Gen. Genet.* 239, 122-128.
117. Wyllie, A.H. (1995). *Current Biology*, 5: 97-104.
118. Zhang, et al. (1988). *Plant Cell Rep.* 7, 379-384.
- 25 119. Zhang, et al. (1991). *Plant Cell* 3, 1155-1165.
120. Zhang, et al. (1988). *Theor Appl Genet* 76, 835-840.

CLAIMS:

1. An isolated polynucleotide encoding a polypeptide which includes the amino acid sequence shown in Figure 2.

2. A polynucleotide according to claim 1 wherein the coding sequence is the coding sequence shown in Figure 2.

3. A polynucleotide according to claim 1 wherein the coding sequence is a mutant, allele, variant or derivative of the coding sequence shown in Figure 2, by way of addition, deletion, substitution and/or insertion of one or more nucleotides.

4. An isolated polynucleotide which on expression in a transgenic plant exerts a negative regulatory effect on a pathogen defence response of the plant, which defence response is pathogen independent and autonomous of the presence of pathogen, the polynucleotide encoding a polypeptide which includes an amino acid sequence which is a mutant, allele, variant or derivative of the Barley Mlo sequence shown in Figure 2, or is a homologue of another species or a mutant, allele, variant or derivative thereof, the amino acid sequence differing from that shown in Figure 2 by way of addition, substitution, deletion and/or insertion of one or more amino acids.

5. A polynucleotide according to claim 4 encoding a polypeptide which includes the amino acid sequence shown in Figure 13.

5 6. A polynucleotide according to claim 5 wherein the coding sequence is that shown in Figure 10.

7. A polynucleotide according to claim 5 wherein the coding sequence is a mutant, allele, variant or derivative of the coding sequence shown in Figure 10, by way of addition, deletion, substitution and/or insertion of one or more nucleotides.

8. A polynucleotide according to claim 4 encoding a polypeptide which includes the amino acid sequence shown in Figure 14.

9. A polynucleotide according to claim 8 wherein the coding sequence is that shown in Figure 11.

10. A polynucleotide according to claim 8 wherein the coding sequence is a mutant, allele, variant or derivative of the coding sequence shown in Figure 11, by way of addition, deletion, substitution and/or insertion of one or more nucleotides.

11. A polynucleotide according to claim 4 encoding a

110

polypeptide which includes the amino acid sequence shown in Figure 15.

12. A polynucleotide according to claim 11 wherein the coding sequence is that shown in Figure 12.

13. A polynucleotide according to claim 11 wherein the coding sequence is a mutant, allele, variant or derivative of the coding sequence shown in Figure 12, by way of addition, deletion, substitution and/or insertion of one or more nucleotides.

14. A polynucleotide according to any preceding claim operably linked to a regulatory sequence for expression.

15. An isolated polynucleotide encoding a polypeptide which on expression in a transgenic plant produces a polypeptide which can stimulate or maintain a defence response of the plant, the encoded polypeptide including an amino acid sequence which is a mutant, allele, variant or derivative of the Barley *Mlo* sequence shown in Figure 2 or of a homologue of another species, the amino acid sequence differing from that shown in Figure 2 by way of addition, substitution, deletion and/or insertion of one or more amino acids.

16. A polynucleotide according to claim 15 which stimulates or maintains said defence response of the plant on homozygous



expression in the plant.

17. A polynucleotide according to claim 15 wherein the amino acid sequence includes an alteration identified in Table 1.

5

18. A polynucleotide according to claim 17 wherein the amino acid sequence is that of Figure 2 including a substitution at residue 240.

10

19. A polynucleotide according to claim 17 wherein the amino acid sequence includes Leucine at residue 240.

15

20. A polynucleotide according to any of claims 15 to 19 operably linked to a regulatory sequence for expression.

20

21. An isolated polynucleotide which has at least about 600 contiguous nucleotides of the nucleotide sequence of any of claims 1 to 13 or complement thereof

25

22. A polynucleotide according to claim 21 operably linked to a regulatory sequence for transcription.

23. An isolated polynucleotide which has at least about 300 contiguous nucleotides of the sequence of any of claims 1 to 13, or complement thereof, operably linked to a regulatory sequence for transcription.

112

24. A polynucleotide according to claim 22 or claim 23 wherein the regulatory sequence includes an inducible promoter.

25. A nucleic acid vector suitable for transformation of a host cell and including a polynucleotide according to any preceding claim.

26. A nucleic acid vector according to claim 25 wherein said host cell is a microbial cell.

27. A nucleic acid vector according to claim 25 wherein said host cell is a plant cell.

28. A host cell containing a heterologous polynucleotide or nucleic acid vector according to any preceding claim.

29. A cell according to claim 28 which is microbial.

30. A cell according to claim 28 which is a plant cell.

31. A cell according to claim 30 having said heterologous polynucleotide incorporated within its genome.

32. A cell according to claim 31 having more than one said polynucleotide per haploid genome.

33. A cell according to any of claims 30 to 32 which is

comprised in a plant.

34. A plant including a cell according to any of claims 30 to 32.

5

35. A plant which is a sexually or asexually propagated off-spring, clone or descendant of a plant according to claim 34, or any part or propagule of said plant, off-spring, clone or descendant.

36. A part or propagule of a plant according to claim 35.

37. A plant according to claim 34 which does not breed true.

38. A method of producing a plant, the method including incorporating a heterologous polynucleotide according to any of claims 1 to 14 into a plant cell and regenerating a plant from said plant cell.

39. A method of producing a plant, the method including incorporating a heterologous polynucleotide according to any of claims 15 to 20 into a plant cell and regenerating a plant from said plant cell.

40. A method of producing a plant, the method including incorporating a heterologous polynucleotide according to any of claims 21 to 24 into a plant cell and regenerating a plant from

said plant cell.

41. A method according to any of claims 38 to 40 including sexually or asexually propagating or growing off-spring or a descendant of said plant.

42. A method of stimulating a defence response in a plant, the method including causing or allowing transcription from a heterologous polynucleotide according to any of claims 1 to 14 within cells of the plant.

43. A method of stimulating a defence response in a plant, the method including causing or allowing transcription from a heterologous polynucleotide according to any of claims 15 to 20 within cells of the plant.

44. A method of stimulating a defence response in a plant, the method including causing or allowing transcription from a heterologous polynucleotide according to any of claims 21 to 24 within cells of the plant.

45. A method of producing a polynucleotide encoding a polypeptide which on expression in a transgenic plant produces a polypeptide which can stimulate or maintain a defence response of the plant, the method including alteration of the nucleotide sequence of a polynucleotide according to any of claims 1 to 14.

115

46. A method according to claim 45 involving site-specific sequence mutation.

47. A method according to claim 45 involving intracellular homologous recombination.

48. A method wherein following alteration of a nucleotide sequence in accordance with the method of claim 45 a polynucleotide including the altered nucleotide sequence is introduced into a host cell.

49. A method according to claim 48 wherein the host cell is a plant cell.

50. A method wherein following introduction of a polynucleotide into a plant cell in accordance with claim 49 a plant is regenerated from the cell or descendants thereof including the altered nucleotide sequence.

51. Use of a polynucleotide according to any of claims 1 to 14 for stimulating a defence response in a plant.

52. Use of a polynucleotide according to any of claims 15 to 20 for stimulating a defence response in a plant.

53. Use of a polynucleotide according to any of claims 21 to 24 for stimulating a defence response in a plant.

54. Use of a polynucleotide according to any of claims 21 to 24 for down-regulation of expression of a gene encoded a polypeptide encoded by a polynucleotide according to any of claims 1 to 14.

5

55. Use of a polynucleotide according to any of claims 1 to 14 in the production of a transgenic plant.

56. Use of a polynucleotide according to any of claims 15 to 20 in the production of a transgenic plant.

57. Use of a polynucleotide according to any of claims 21 to 24 in the production of a transgenic plant.

58. A method of determining the presence of a pathogen resistance or susceptibility allele in a plant or plant cell, the method including analysing a sample from the plant or plant cell by:

(a) comparing the sequence of nucleic acid in the sample with all or part of the nucleotide sequence shown in Figure 7 to determine whether the sample from the patient contains a mutation;

(b) determining the presence in the sample of a polypeptide including the amino acid sequence shown in Figure 7 or a fragment thereof and, if present, determining whether the polypeptide is full length, and/or is mutated, and/or is expressed at the normal level;

(c) performing DNA fingerprinting to compare the restriction pattern produced when a restriction enzyme cuts nucleic acid in the sample with the restriction pattern obtained from the nucleotide sequence shown in Figure 7 or from a known mutant, allele or variant thereof;

(d) contacting the sample with a specific binding member capable of binding to nucleic acid including the nucleotide sequence as set out in Figure 7 or a fragment thereof, or a mutant, allele or variant thereof, the specific binding member including nucleic acid hybridisable with the sequence of Figure 7 or a polypeptide including a binding domain with specificity for nucleic acid including the sequence of Figure 7 or the polypeptide encoded by it, or a mutated form thereof, and determining binding of the specific binding member;

(e) performing PCR involving one or more primers based on the nucleotide sequence shown in Figure 7 to screen the sample for nucleic acid including the nucleotide sequence of Figure 7 or a mutant, allele or variant thereof.

59. A method of determining the presence of target nucleic acid in a plant or plant cell, the method including contacting a nucleic acid molecule which includes the nucleotide sequence shown in Figure 7 or an oligonucleotide fragment thereof with nucleic acid in a sample from the plant or plant cell and assessing hybridisation of said nucleic acid molecule with nucleic acid in the sample.

60. A method according to claim 59 which involves amplification of nucleic acid to which said nucleic acid molecule hybridises.

5 61. A method according to claim 59 or claim 60 wherein said nucleic acid molecule includes an alteration in sequence compared with the nucleotide sequence shown in Figure 7 or corresponding fragment thereof.

10 62. A method according to claim 61 wherein said alteration is selected from those shown in Table 1.

15 63. An assay method for identifying a compound able to bind the polypeptide encoded by the polynucleotide of any of claims 1 to 14 or any of claims 15 to 20, the method including:

(a) bringing into contact said polypeptide or a fragment thereof, and a test compound; and

(b) determining interaction or binding between said polypeptide or fragment thereof and the test compound.

20 64. An assay method according to claim 63 wherein a compound is identified which is able to bind the polypeptide for which the amino acid sequence is shown in Figure 2.

25 65. An assay method for identifying a compound able to stimulate a defence response in a plant by interaction with the polypeptide encoded by the polynucleotide of any of claims 1 to



14 or any of claims 15 to 20, the method including:

(a) contacting a plant or plant part with a test compound and determining stimulation of a defence response; and

(b) bringing into contact said polypeptide or a fragment thereof with a test compound and determining interaction or binding between said polypeptide or a fragment thereof and the test compound;

step (b) being performed with a test compound which tests positive in step (a), or step (a) being performed with a test compound which tests positive in step (b), or steps (a) and (b) being performed in parallel.

66. An assay method according to claim 65 wherein stimulation of a defence response is determined by monitoring pathogen growth and/or viability on the plant or plant part.

67. An assay method according to claim 65 or claim 66 wherein a compound is identified which is able to bind the polypeptide for which the amino acid sequence is shown in Figure 2.

68. An assay method according to any of claims 65 to 67 wherein a compound is identified which is able to stimulate resistance to powdery mildew in barley.

69. A method which includes following identification of a compound as being able to stimulate a defence response in a plant in accordance with any of claims 65 to 68 formulation of

120

the compound, or optionally if the compound is peptidyl nucleic acid encoding it, into a composition including at least one additional component.

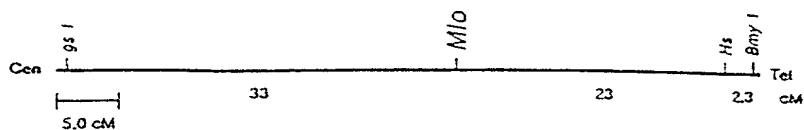
5 70. A method which includes following identification of a compound as being able to stimulate a defence response in a plant in accordance with any of claims 56 to 58 application of the compound, or optionally if the compound is peptidyl nucleic acid encoding it, to a plant.

10 71. Use of a polypeptide encoded by a polynucleotide according to any of claims 1 to 14, in screening for compounds able to stimulate a defence response in a plant.

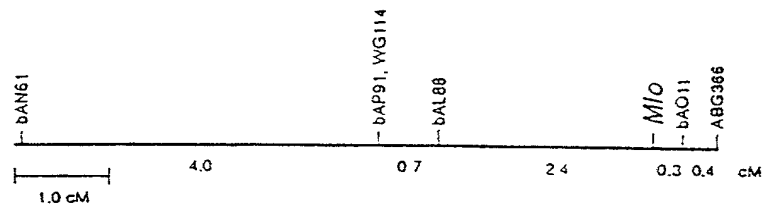
15 72. Use of a polypeptide encoded by a polynucleotide according to any of claims 15 to 20, in screening for compounds able to stimulate a defence response in a plant.

20 73. A compound able to stimulate a defence response in a plant identified by a method according to any of claims 63 to 68.

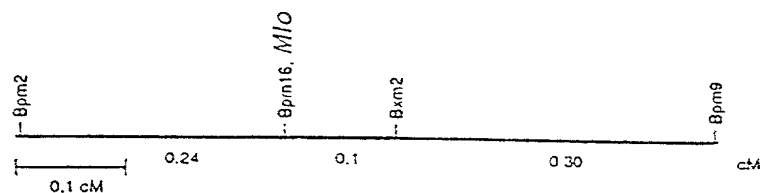
1/28

Morphological  
marker map

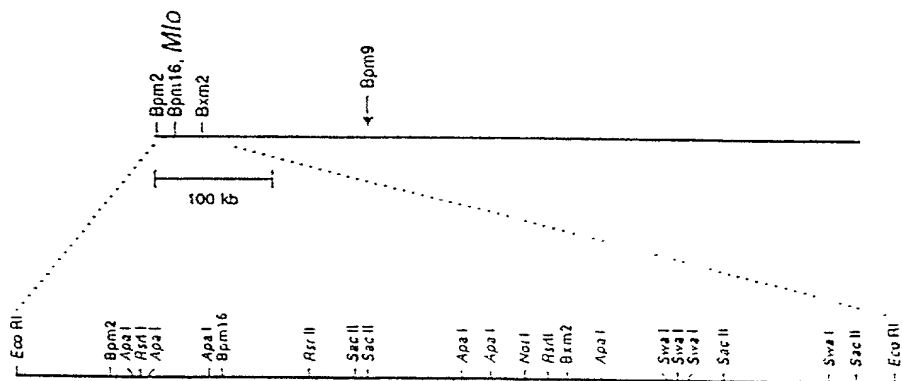
RFLP map

(CarlsbergII *Mlo* x  
Grannenlose Zweizeilige *mlo-11*)

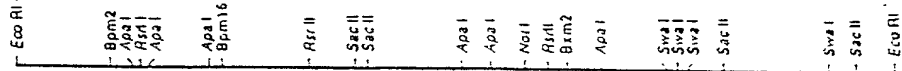
AFLP map

(Ingrid *Mlo* x BC Ingrid *mlo-3*)

YAC YHV303-A6



BAC F15



Sequence contigs

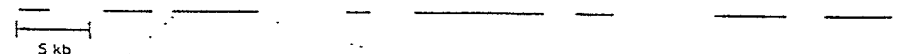
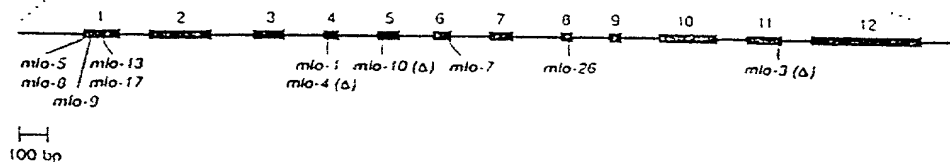
*Mlo* gene structure

Figure 1

2/28

M S D K K G V P A R E L P E T P S W A V  
 ATGTCGGACAAAAAGGGGTGCCGGCGCGGGAGCTGCCGGAGACGCCGTCGTGGGCGGTG 60  
 A V V F A A M V L V S V L M E H G L H K  
 GCGGTGGTCTTCGCCGCCATGGTGCTCGTGTCCTCATGGAACACGGCCTCCACAAG 120  
 L G H W F Q H R H K K A L W E A L E K M  
 CTCGGCCATTGGTTCCAGCACCGGCACAAGAAGGCCCTGTGGGAGGCGCTGGAGAAGATG 180  
 K A E L M L V G F I S L L L I V T Q D P  
 AAGGCGGAGCTCATGCTGGTGGGCTTCATATCCCTGCTCCTCATCGTCACGCAGGACCCC 240  
 I I A K I C I S E D A A D V M W P C K R  
 ATCATCGCCAAGATATGCATCTCCGAGGATGCCGCCGACGTCATGTGGCCCTGCAAGCGC 300  
 G T E G R K P S K Y V D Y C P E G K V A  
 GGCACCGAGGGCCGCAAGCCCAGCAAGTACGTTGACTACTGCCCGGAGGGCAAGGTGGCG 360  
 L M S T G S L H Q L H V F I F V L A V F  
 CTCATGTCCACGGGCAGCTTGCAACAGCTGCACGTCTTCATCTTCTGTGCTCGCGGTCTTC 420  
 H V T Y S V I T I A L S R L K M R T W K  
 CATGTACCTACAGCGTCATCACCATAGCTCTAAGCCGTCTCAAAATGAGAACATGGAAG 480  
 K W E T E T T S L E Y Q F A N D P A R F  
 AAATGGGAGACAGAGACCACCTCCTTGAATACCAAGTTCGCAAATGATCCTGCACGGTTC 540  
 R F T H Q T S F V K R H L G L S S T P G  
 CGGTTACGCACCAGACGTCGTTCGTGAAGCGCCACCTGGGCCTCTCCAGCACCCCTGGC 600  
 I R W V V A F F R Q F F R S V T K V D Y  
 ATCAGATGGGTGGTGGCCTTCTTCAGGCAGTTCCTCAGGTCAGTCACCAAGGTGGACTAC 660  
 L T L R A G F I N A H L S Q N S K F D F  
 CTGACCTTGAGGGCAGGCTTCATCAACGCGCATTTGTGCGAAAACAGCAAGTTCGACTTC 720  
 H K Y I K R S M E D D F K V V V G I S L  
 CACAAGTACATCAAGAGGTCGATGGAGGACGACTTCAAGGTCGTCGTGCGCATCAGCCTC 780  
 P L W G V A I L T L F L D I N G V G T L  
 CCGCTGTGGGGTGTGGCGATCCTCACCTCTTCTTGACATCAATGGGGTGGCAGCTC 840  
 I W I S F I P L V I L L C V G T K L E M  
 ATCTGGATTCTTTCATCCCTCTCGTGATCCTCTTGTGTGTTGGAACCAAGCTGGAGATG 900  
 I I M E M A L E I Q D R A S V I K G A P  
 ATCATCATGGAGATGGCCCTGGAGATCCAGGACCGGGCGAGCGTCATCAAGGGGGCCCCC 960  
 V V E P S N K F F W F H R P D W V L F F  
 GTGGTCGAGCCAGCAACAAGTTCTTCTGGTTCACCGCCCCGACTGGGTCTCTTCTTC 1020  
 I H L T L F Q N A F Q M A H F V W T V A  
 ATACACCTGACGTGTTCAGAACCGGTTTCAGATGGCGCATTTTGTGTGGACAGTGGCC 1080  
 T P G L K K C Y H T Q I G L S I M K V V  
 ACGCCCGGCTTGAAGAAATGCTACCACACGCAGATCGGGCTGAGCATCATGAAGGTGGTG 1140  
 V G L A L Q F L C S Y M T F P L Y A L V  
 GTGGGGCTAGCTCTCCAGTTCCTCTGCAGCTATATGACCTTCCCCCTCTACGCGCTCGTC 1200  
 T Q M G S N M K R S I F D E Q T S K A L  
 ACACAGATGGGATCAACATGAAGAGGTCCATCTTCGACGAGCAGACGTCCAAGGCGCTC 1260  
 T N W R N T A K E K K K V R D T D M L M  
 ACCAACTGGCGGAACACGGCCAAGGAGAAGAAGAAAGTCCGAGACACGGACATGCTGATG 1320  
 A Q M I G D A T P S R G S S P M P S R G  
 GCTCAGATGATCGGCGACGCAACACCGAGCCGAGGCTCGTCCCGATGCCGAGCCGGGGC 1380

Figure 2

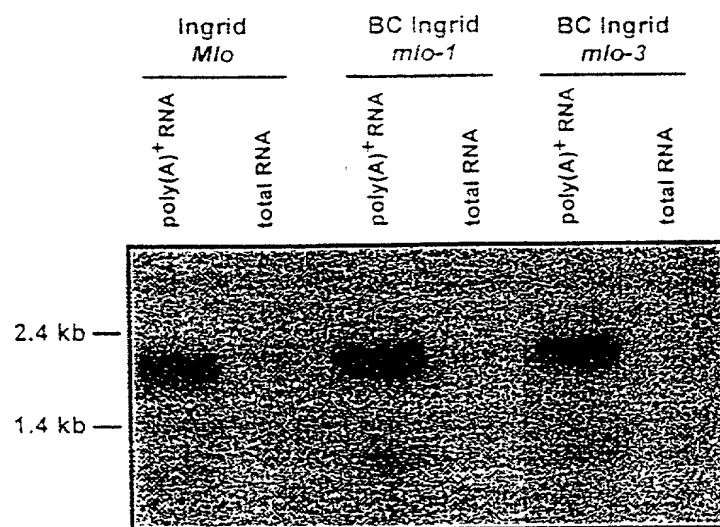
3/28

Figure 2 (Continued)

S S P V H L L H K G M G R S D D P Q S A 1440  
TCATCACCCGTGCACCTGCTTCACAAGGGCATGGGGCGGTCGGACGACCCCCAGAGCGCG  
P T S P R T Q Q E A R D M Y P V V V A H 1500  
CCCACCTCGCCAAGGACCCAGCAGGAGGCTAGGGACATGTACCCGGTTGTGGTGGCGCAC  
P V H R L N P N D R R R S A S S S A L E 1560  
CCGGTGCACAGACTAAATCCTAACGACAGGAGGAGGTCCGCCTCGTCGTCGGCCCTCGAA  
A D I P S A D F S F S Q G \*  
GCCGACATCCCCAGTGCAGATTTTTCCTTCAGCCAGGGATGA 1602

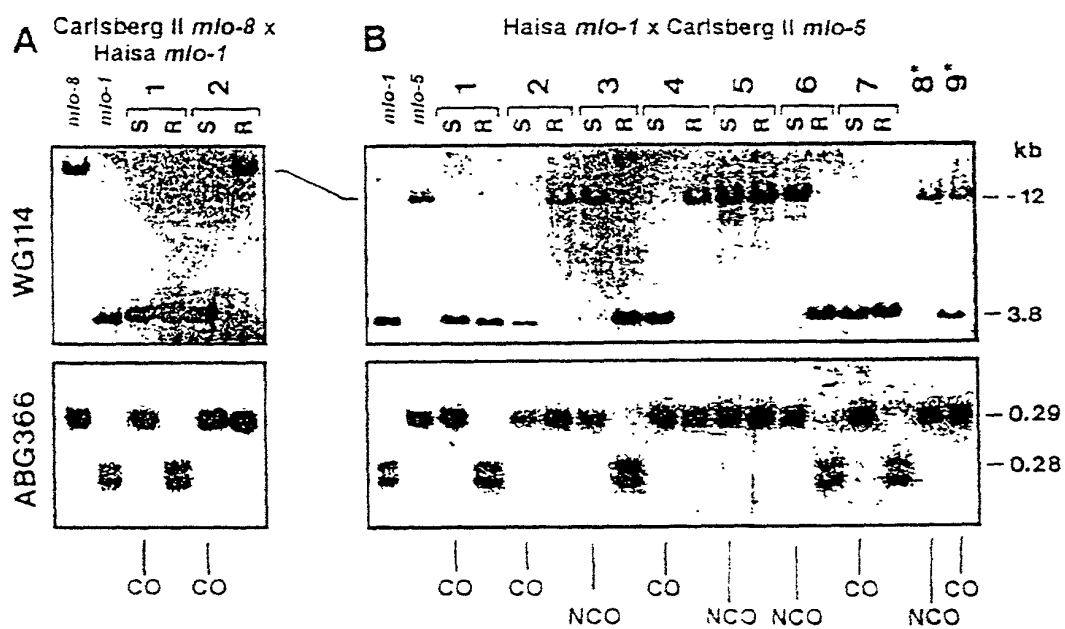
09/23/98 11:20:00

4/29

Figure 3

5/28

Figure 4



292 GCGGAGCTCATGCTGGTGGGCTTCATATCCCTGCTCCTCATCGTCAGCA 341  
 || : ||| |||| ||||| ||||| : ||||| ||| ||| |||  
 80 GCANAGCTGATGCTGCTGGGCTTCATNTCCCTGCTTCTCACCCTGGCACA 129  
 342 GGACCCCATCATCGCCAAGATATGCATCTCCGAGGATGCCGCCGACGTCA 391  
 || || ||| |||| : || ||||| || || || ||| ||| |||  
 130 GCGGCC...CATCTCCAANATCTGCATCCCCAAGTCGGCTGCCAACATCT 176  
 392 TGTGGCCCTGCAAGCGCGGCACCGAGGGCCGC..AAGCCCAGCAAGTACGT 440  
 ||| ||| ||||| ||| : ||| ||| ||| : ||| ||| |||  
 177 TGTGCGCTGCAAGGCAGGCCNAGATGCCATCGAAGAANAAGCAGCAAGT 226  
 441 TGACTACTGCCCCGAGGTGAGCAGCAGAGCCCGGACCAGCAGCTTCACGA 490  
 || : ||| || || || || || || || ||| : |||  
 227 GGTNCNCNGTCC..TTGGCCGGCGCCGGCGGGGACTACTGCTCNAAT 275  
 491 TGATGAAGAAATCAATACC.....GAACTTTTCTTGTTTTCT 528  
 | || | || : || | || | || : : :  
 276 TCGATGTGAGAATAACNCCAGCTGCCGGCAAGCACAACCTCGATNCNATN 325  
 529 TCTGATTGTCTGCTTGGCTTGGCTTAATTGGTGTGTGTGTGTGTGTTGC 578  
 ||: ||| | ||||| ||||| ||| ||| ||| |||  
 326 ACTNATT.....TAACTATAATTGATTTTCTTGGGTTTCTGC 364  
 579 AGGGCAAGGTGGCGCTCATGTCCACGGGCAGCTTGCACCAAGCTGCACGT 628  
 ||||| ||||| |||| | ||| ||||| ||||| |||||  
 365 AGGGCAAGGTGGCGCTGATGTCCGCAAGAGCATGCACCAGCTGCACATT 414  
 629 TTCATCTTCGTGCTCGCGGTCTTCCATGTACCTACAGCGTCATCACCAT 678  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 415 TTCATCTTCGTGCTCGCCGTGTTCCATGTTACCTACTGCATCATCACCAT 464  
 579 AGCTCTAAGCCGTCTCAAAGTGAGCCTTTGCTTCT....TCTTCTTCTT 723  
 ||| ||||| ||||| ||| ||||| ||| ||| |||  
 465 GGGTTTAGGGCGCCTCAAAGTGAGTTTGTCTGTTCTGTCCCTCATGCACAT 514  
 724 CTTTTACC.....GCACGTCTGTCTGTACGGCGTACCTACCTGTTCA 765  
 |||| | ||| : ||| ||| ||| ||| ||| |||  
 515 GTTTTCTCTAGTTCTAGCAANATTGTCAGTCCTTCAAATGGATTGTTTCG 564  
 766 TCAGGCTTGAGTAAACTGTTCATAATCTGC.....TCCGGCATAA 807  
 || || || || || || || || || || || || || || || ||  
 565 ACA.....AGAAACCAATTATTAATTGCCAGTTAAATATATAATAA 608  
 808 TCCTCTCCTCCTG...CAGATGAGAACATGGAAGAAATGGGAGACAGAG 853  
 || || || || || || || || || || || || || || || ||  
 609 TTGATCTTTCTTGGTTTTAGATGAAGAAATGGAAGAAGTGGGAGTCACAG 658  
 854 ACCACCTCCTTGGAATACCAGTTTCGCAATGGTCAGGATCCCCCACTCTG 903  
 |||| ||| |||| || ||||| |||| ||| ||| |||  
 659 ACCAACTCATTGGAGTATCAGTTCGCAATCGGTAGTG.....AATTAA 701  
 904 CAATCTCCC...CTTCTTCGAAACCAACC...TGATGATCCATTATAA 946  
 ||||| || || || || || || || || || || || || || || || ||  
 702 GAATCTCCCTAACTATTTCAATTCAGAACCTTTATGATAATGTCTTGAA 751  
 947 GACGCAGGCACGATCAGAGTGAGTGAAGTATGTATGTTTCTTTTGTG 996  
 ||| || ||| || || || || || || || || || || || || || || || ||  
 752 GAGGAGGAGCAAATCAG..CTGAAAAATATGATCGA..... 785  
 997 TCCTTTCAGATCCTGCACGGTTCCGGTTCACGCACCAGACGTCGTTTCGTG 1046  
 ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 786 TCCATGCAGATCCTTCACGATTGAGTTTACGCATCAGACGTCGTTTCGTG 835



7128

Variable	Mean	SD	Min	Max
Age	34.5	10.2	21	55
Gender	0.5	0.5	0	1
Marital status	0.6	0.5	0	1
Education	12.5	1.5	9	16
Income	1500	500	500	3000
Health status	0.8	0.2	0	1
Stress level	3.5	1.5	1	5
Life satisfaction	4.2	1.0	1	5
Work satisfaction	3.8	1.2	1	5
Family satisfaction	4.0	1.1	1	5
Community satisfaction	3.9	1.3	1	5
Overall satisfaction	4.1	1.0	1	5

FIGURE 5 cont'd

1846 CGAAAGCAACAGCAAACCCAATT.....TGATCGCAAT 1878  
 | | | | | | | | | | | | | | | | | | | | | |  
 1619 CCAAAGAAACAGTTAACATAATTAATTATGCTTTAATTTATCATGAAAAT 1668  
 1879 GGAAACCCACACCTAATATTAACCTCAAAATGTCAATTGTCGGTGGCTCTT 1928  
 | | | | | | | | | | | | | | | | | | | | | |  
 1669 TAATATGATCATATAACTAATGAACAAACATTCA..TGTGAATGCCACCG 1716  
 1929 CCTCAACAGATCCTCTTGTGTGTGGAAACCAAGCTGGAGATGATCATCAT 1978  
 | | | | | | | | | | | | | | | | | | | | | |  
 1717 TTGTCTCAGATCGTCTTGTAGTTGGGACCAAGCTAGAGATGGTGATCAT 1766  
 1979 GGAGATGGCCCTGGAGATCCAGGACCGGGCGAGCGTCATCAAGGGGGGCC 2028  
 | | | | | | | | | | | | | | | | | | | | | |  
 1767 GGAGATGGCCCAAGAGATACAGGACAGGGCCACTGTGATCCAGGGAGCAC 1816  
 2029 CCGTGGTCGAGCCCAAGCAACAAGTTCTTCTGGTTCCACCGCCCGACTGG 2078  
 | | | | | | | | | | | | | | | | | | | | | |  
 1817 CTATGGTTGAACCAAGCAACAAGTACTTCTGGTTCAACCGCCCTGACTGG 1866  
 2079 GTCCTCTTCTTCATACACCTGACGTTGTT..... 2107  
 | | | | | | | | | | | | | | | | | | | | | |  
 1867 GTCTTGTTCTTCATACACCTGACACTCTTCCCATGTACATGTTTAAACC 1916  
 :  
 :  
 :  
 2108 .....CCAGAACGC.GTTTCAGATGGCGCATTTTG 2136  
 | | | | | | | | | | | | | | | | | | | | | |  
 2017 GACGGACGGATCGATCATCACCAGAACGCATTTTCAGATGGCGCATTTTCG 2066  
 2137 TGTGGACAGTG....GTACGCCAC.....CGATGA<sup>^</sup>ACTGTGCAGTT 2173  
 | | | | | | | | | | | | | | | | | | | | | |  
 2067 TATGGACTATGGTGTGTATGCTACTTGCTTAGTTGTTGCCATTATCAGTT 2116  
 2174 .....AACATGGGTGTCA...AGGCACCGAGTGCCGCTGATGA.... 2208  
 | | | | | | | | | | | | | | | | | | | | | |  
 2117 CTTAAGCAAATTAAGTGTGATGCATGCACTGA.....CTAATGAGACAA 2160  
 2209 .....ACTGCTCTGACGGAGATTTACTTGTGTTGT.....AGGCC 2243  
 | | | | | | | | | | | | | | | | | | | | | |  
 2161 AAAATGACACAGCTTGTTTCATCGATCTGGTTGTTTGTGTGTGACAGGCA 2210  
 2244 ACGCCCGGCTTGAAGAAATGCTACCACACGCAGATCGGGCTGAGCATCAT 2293  
 | | | | | | | | | | | | | | | | | | | | | |  
 2211 ACACCTGGTCTGAAGAAATGCTTCCATGAAAATATTTGGCTGAGCATCGT 2260  
 2294 GAAGGTGGTGGTGGGGCTAGCTCTCCAGTTCCTCTGCAGCTATATGACCT 2343  
 | | | | | | | | | | | | | | | | | | | | | |  
 2261 GGAAGTCATTGTGGGGATCTCTCTTCAGGTGCTATGCAGCTACATCACCT 2310  
 2344 TCCCCCTCTACGCGCTCGTCACACAGGTAATAAAACCGTCCAGGAA 2389  
 | | | | | | | | | | | | | | | | | | | | | |  
 2311 TCCCGCTCTACGCGCTCGTCACACAGGTGAACAAGCCATTACAAA 2356

FIGURE 6

Parameter	Value	Unit
Temperature	25.0	°C
Pressure	1.0	atm
Flow rate	1.0	L/min
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	nm
Scan rate	1.0	nm/min
Integration time	1.0	s
Resolution	0.5	nm
Detector	Photodiode array	
Injection volume	10	μL
Mobile phase	Water/Acetonitrile (90/10)	
Column	C18, 150 × 4.6 mm, 5 μm	
Flow rate	1.0	mL/min
Temperature	30.0	°C
Wavelength	254	

FIGURE 6 cont'd

Figure 1 consists of 11 pie charts, labeled (a) through (k), arranged vertically. Each chart represents a different category of land use or vegetation cover. The data for each chart is as follows:

- (a) Land use: agricultural (45%), forest (35%), urban (15%), water (5%).
- (b) Vegetation cover: forest (60%), agricultural (25%), urban (10%), water (5%).
- (c) Land use: agricultural (40%), forest (30%), urban (20%), water (10%).
- (d) Vegetation cover: forest (55%), agricultural (20%), urban (15%), water (10%).
- (e) Land use: agricultural (35%), forest (25%), urban (25%), water (15%).
- (f) Vegetation cover: forest (50%), agricultural (15%), urban (20%), water (15%).
- (g) Land use: agricultural (30%), forest (20%), urban (30%), water (20%).
- (h) Vegetation cover: forest (45%), agricultural (10%), urban (25%), water (20%).
- (i) Land use: agricultural (25%), forest (15%), urban (35%), water (25%).
- (j) Vegetation cover: forest (40%), agricultural (5%), urban (30%), water (25%).
- (k) Land use: agricultural (20%), forest (10%), urban (40%), water (30%).

FIGURE 6 cont'd

[illegible]

12/28

FIGURE 6 cont'd

```

2632 CGGGGCTCATCACCCGTGCACCGTTCACAAAGGGCATGGGGCGGGTCGA 2681
      |||||
2108 .....GTGCACCTGCTCCACAAGGCCGGGGCGCGGTCCGA 2142
2682 CGACCCCCAGAGCGCGCCCACCTCGCCAAGGACCCAGCAGGAGGCTAGGG 2731
      |||||
2143 CGACCCCCAGAGCGTGCCGGCGTCCCCGAGGGCCGAGAAGGAAGGCGGCG 2192
2732 ACATGTACC CGGTGTGTGGTGGCGCACCCGGTGCACAGACTAAATCCTAAC 2781
      |          |||   |||   |||   |||   |||   |||   |||
2193 GC.....GTGCAGCATCCGGCGCGCAAGGTACCTCCTTGT 2227
2782 GACAGGAGGAGGTCCGCCTCGTCGTCGGCCCTCGAAGCCGACATCCCCAG 2831
      |||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
2228 GACGGGTGGAGGTCCGCCTCGTCGCGGCGCTCGACGCTCACATCCCCGG 2277
2832 TGCAGATTTTTCCTTCAGC.....CAGGGATGAGACAAGTTTCTG 2871
      |||||           |||   |||   |||   |||   |||   |||
2278 TGCAGATTTTGGCTTTCAGCACGCAACGTTGACCGATCAGACAAGTTCCTT 2327
2872 TATT 2875
      |  ||
2328 TTTT 2331

```

Figure 7

13/28

helix I

helix II

helix III

helix IV

helix V

helix VI

GGCTGCTCGGCCAGCAACACAGACACAGCAGCGTACCTGCGT  
 ACGTAGCGTGCGCTTCTTTTTCCTTTCGCTCTCTTGCTGCTCGGGCGGCCAGG  
 TCGATAGCCGGCCACGGCCAGGCACCTCGCGGTGCGTGCCTGCATCTGCGTGTGCGTA  
 CCTGGTAGAGCGGCCCTCTGCTTGTCTCGGGCAAGGAAGGAGTTGCGGCGGTGCGACCG  
 M S D K K G V P A R E L P E T P S W R A V 20  
 ATGTGGGACAAAAAGGGTGC CGGCGCGGAGCTGCGGAGACGCGTCTGTGGCGGTG 60  
A V V V F A A M V L V S V L M E H G L H K 40  
 GCGGTGCTTTCGCGCCATGGTGTCTGCTGCTCTCATGSAACACGGCTCCACAAG 120  
 L G H W F Q H R H K K A L W E A L E K M 60  
 CTCGGCCATTGGTTCACGACCGGCACAAGAAGGCCCTGTGGGAGGCGCTGGAGAAGATG 180  
 K A E L M L V G F I S L L L I V A T Q R D A D 80  
 AAGGCGGAGCTCATGCTGGTGGGCTTCATATCCCTGCTCTCATGTCACGCAGGACCC 240  
I R A A K I C I S E D A A D V M W P C K R 100  
 ATCATCGCAAGATATGCATCTCCAGGATGCCCGGACGTATGTGGCCCTGCAAGCGC 300  
 G T E G R K P S K Y V D Y C P E G K V A 120  
 GGCACCGAGGGCCSCAAGCCAGCAAGTAGCTTGACTACTGCCGAGGGCAAGGTGGC 360  
 L M S T G S L H Q L H V F I F V L A V A F 140  
 CTCATGTCCACGGGCGAGCTTGACACAGCTGCAGTCTTGATCTCGTGCTCGGGTCTTC 420  
H V T Y S V I T I A L S R L K M R T W K 160  
 CATGTCACTACAGCGTCAACCATAGCTCTAAGCCGTCTCAAAATGAGAATGGAAG 480  
 K W E T E T T S L E Y Q F A N D P A R F 180  
 AAATGGGAGACAGAGACCGCTCTTGAATACCACTTGCAGATGATCCTGCACGGTTC 540  
 R F T H Q T S F V K R H L G L S S T P G 200  
 CGGTTACGCAAGACGTCTGTTGTAAGCGCCAGCTGGGCTCTCCAGCACCCCTGGC 600  
 I R W V V A F F R Q F F R S V T K V D Y 220  
 ATCAGATGGGTGGCTTCTTCAGGCAGTCTTCAGGTCACTACCAAGGTGGACTAC 660  
 L T L R A G F I N A H L S Q N S K F D F 240  
 CTGACCTTGAGGGCAGGCTTCATCAACCGCATTTGTCGCAAAACAGCAAGTTCGACTTC 720  
 H K Y I K R S M E D D F K V V V G I S L 260  
 CACAAGTACATCAAGAGGTGATGGAGGACGACTTCAAGGTGCTGCTCGGCATCAGGCTC 780  
P L W G V A I L T L F L D I N G V G T A L 280  
 CCGCTGTGGGGTGTTGGCGATCTCAGCCCTCTTCCTTGACATCAATGGGTGGCAGCTC 840  
I W I S F I P L V I L L C V G T X L E M 300  
 ATCTGGATTCTTTCATCGCTCTCGTGATCCTCTGTGTGTTGAACCAAGCTGGAGATG 900  
 I I M E M A L E I Q D R A S V I K G A P 320  
 ATCATCATGGAGATGGCCCTGGAGATCCAGGACCGGGCGAGCTCATCAAGGGGGCCCC 960  
 V V E P S N K F F W F H R P D W V L F F 340  
 GTGGTCGAGCCAGCAACAGTCTCTTGTGTTCCACCGCCCGACTGGGTCTCTCTTC 1020  
 I H L T L F Q N A F Q M A H F V W T V A 360  
 ATACACCTGACGTGTTCAGAACCGGTTTCAGATGGCGCATTTGTGTGGACAGTGGCC 1080  
 T P G L K K C Y H T Q I G L S I M K V R V 380  
 ACGCCGGGCTGAAGAAATGCTACACAGCGAGATCGGGCTGAGCATCATGAAGGTGGTG 1140  
V G L A D Q F L C S Y M T F P L Y A L A V 400  
 GTGGGCTAGCTCTCCAGTTCCTCTGCAGCTATATGACCTTCCCTCTACGCGCTCGTC 1200  
T Q M G S N M K R S I F D E Q T S K A L 420  
 ACACAGATGGGATCAACATGAAGAGGTCCATCTTCGACGAGCAGAGTCCAAGGCGCTC 1260  
 T N W R N T A K E K K K V R D T D M L M 440  
 ACCAACTGGCGGAACAGGCCAAGGAGAAGAAGATCCGAGACAGGACATGCTGATG 1320  
 A Q M I G D A T P S R G S S P M P S R G 460  
 GCTCAGATGATCGGCGACCAACACCGAGCGAGGCTCTCGCGATGCGGAGCGGGG 1380  
 S S P V H L L H K G M G R S D D P Q S A 480  
 TCATCACCCGTGCACCTGCTTCACAAGGGCATGGGCGGTGCGAGACCCCCAGAGCGG 1440  
 P T S P R T Q Q E A R D M Y P V V V A H 500  
 CCCACCTCGCAAGGACCCAGCAGGAGGCTAGGACATGTACCGGTTGTGGTGGCGCAC 1500  
 P V H R L N P N D R R R S A S S S A L E 520  
 CCGGTGCACAGACTAAATCCTAACGACAGGAGGAGTCCGCTCTGCTGCGGCCCTCGAA 1560  
 A D I P S A D F S F S Q G  
 GCGGACATCCCGAGTCAGATTTTCTTCAGCCAGGATGAGACAAGTTTCTGTATTCA  
 TGTAGTCCCAATGTATAGCCAACATAGGATGTGATGATTCGTACAATAAGAAATACAAT  
 TTTTACTGAGTC

09337-12300

## Figure 8

14/28

1 GAATTCAATT AAGGACAACA ACGGATGATA GGCTTAAGCT AGAGAGGATT  
 51 CATATCGATT AATTAAGTGT ACTTAAGTTG AGGTAAAAC CTATCGATTG  
 101 CTTTGGACAC CGGCTCTCCC ATGATCTGCC AAGTTGAGCC GGCCTACCTA  
 151 ATTTTCTTCG AAAGCACACA ACAAACGAAG GTAACCACTA ATCTAGACAC  
 201 CACGCCTAAG TTATCAATTA CTACTCTAGT CTCGCGTAGA AACTTCATTG  
 251 TTTATGGAGA GTGCTAGTAC TAGAGTACTT AATATAATAG TAAGCGACAA  
 301 ACCCACGACG ATGAGAATGT ACCTCACTTA CGTAGTCAAT TAAGTCGAAA  
 351 AGGAAATCTT GAACACTTAC TTTATTAAAG AAGTATTCCT CGAGGTACAG  
 401 GAGAGGAGAG CACGCCAATA ACTCCAGCAC TCCTCCGAAA CCTTTCTCAC  
 451 TCTCTACCCT TTTTCTCCAC ACAACTAAAA TGATGTCTAA TGTATGAAAG  
 501 TGAGTTGTAC TCTATTTTGT TGTGTGTTTG GAAGTGAAAT TAGCTCATCC  
 551 TTTTATAGCA ACTTAATGGT CGGTTGTAGG TTGGTAATTA AGTCGGTAAA  
 601 CACTCACAAC CACCATCGTC AACCAATAGG AGATCGCCAC ATGATCGAAA  
 651 GCTGACAGTT AGGGGTGCCA ACCCTGTTTT GTCCGAACCA AGCAAACAAC  
 701 CTCTATCTAG GACCTCTCTT CTATGCTCTGA CAAGTCGGCC CATATGCCGG  
 751 TGCACATGAG ATTAAGTCAA TTTCAGTCGT TTTGGACTGT CATGTGGGCC  
 801 CTTCCAATCC TTGTGCTCCC ATATGATTGG TCGAAAGTAC ATTTAATTC  
 851 TGGGTGAGTG CTAGAACTAA TATGATAGAT GTGCTCCGGC TCCTGGGAAA  
 901 GAGGCCACTT GACATACTTG GGGTAGTGCC CCAAGGGTAT TCCCTATCGC  
 951 TTTTTCATAA TTTTCTCTCT CCAAAATCGG ACGGAAACAA TAAAAAGAG  
 1001 AGGCGATGTT CATCGGCAAA TATCTATTTT TTTGATAGTG TCTTCCCTTA  
 1051 AAACCTTGATT TTTGCGAAGA CTTCCGGCTA AAACCATGAA ATCAGAGTTC  
 1101 CTTGTAAACA ATTTAATTTG CCTAAATACA AAAAAGATCG AATGGAGATA  
 1151 GCATTAAACT TGCTCCATAC GAATCATATT AGTTGGACCG TAACTCATAG  
 1201 AAAAAGTTGC AAGTTGGTTG ACCTATCAAC CCTCTTATGT TGACCGTAAA  
 1251 CCTGTTATGC ATTAAGGATT AAGTACCGGC AGATCGTCAC TACTCACGAA  
 1301 TGCACAAATT TCCGGTAACG TAGGATGGGA TGAGTTGGTC AGAAACGGGT  
 1351 CACCACGTCG CCCAACCTGC CGCGATCGAG CCATTGGCCG GCGATGCACG  
 1401 CGCTTTGACA CAGCCGCCCC CGGCCCCCGG GCGCGCCCCG GTTTTAAATA  
 1451 AAAACCGGCC GCCCCCTGTC AAAGGTGTCA AAGTGTCAAG TGCATCAGAG  
 1501 CTAAGCTAGC GGTCAACCCAG TCAGCTCACC CCGAGACGCA CCAGGGGATC  
 1551 TATCGGATCA TGGCAGGTGG GAGATCGGGA TCGCGGGAGT TGCCGGAGAC  
 1601 GCCGACGTGG GCGGTGGCCG TCGTCTGCGC CGTCCTCGTG CTCGCTCCG  
 1651 CCGCCATGGA GCACGGCCTC CACAACCTCA GCCATGTACG CGCGCGCGCA  
 1701 CGCGGTGTGC TCATCTCTCG AGTTAATTTG GTTGTGTTG TTGTTGTGTT  
 1751 CTTGTGACAT CTCAATTAAC ATCCGATCGT GGTGATCGA TCGCCCTGTG  
 1801 GTGGCGATAC TGCTTGCAAT GCAGTGCTC CGTAGGCGGC AGAAGAACGC  
 1851 CATGCGCGAC GCCCTCGACA AGATCAAAGC AGGTCAACCT CAGCCTCAGC

097637744300



15/28

## FIGURE 8 cont'd

1901 TCACCCTCAG CCTCCATCTC TAAATATTTG ACGCCGTTGA CTTTTTTTAA  
1951 TATGTTTGAC CATTCGTCCT ATTTAAAAA TTTAAGTAAT TATTAATTCT  
2001 TTTTCTACCA TTTGATTCAT TGCTAAATAT ACTATTATGT ATACATATAG  
2051 TTTTACATAT TTCACTAAAG TTTTTAAATA AGACGAATGG TCAAACATGT  
2101 TTAAAAAAGT CAACGGCGTC AAACATTTAG GAAGAAGAGA ATATTATATT  
2151 GCTGCTCCCC TCTAGCCACT TTGCTGCCTC CCTCGTCATT TTTTCAAGTA  
2201 TTTTACGCAA GACTGGTCCT CCAAATCAAA CGTCACAAAT AAGCCATTTA  
2251 TAGTTTCCCT TCGCTTTTFA AGGGGGACTA CTTGTATTTA ATCATGGAGG  
2301 AAACACCAG TCGGATGTCC GATTACTTAA AAAAAAATTC GGGGGACTAA  
2351 TTTTTTTGGC TGATCATCGG TGAAATATTA GGTATATAT GTTGAAAAA  
2401 AATCAGCCAC AAACAATGAA ATATTTTGTG AAACACATAT TAGACACGTT  
2451 GAAACGTATC ATTGTTACGT ATAAAAACATC GAATGTTAAC AGATTAAAAAC  
2501 ATATGTTTTT TTTTAATCAG AATATAATCA TGCGATATAT TATTGTAAAG  
2551 ATATAATTAC AACGAATACA ACAGTGCAT CGGATTATAT ATATATTAGT  
2601 AGTTTAAGAG AAAAATCATT TTGAAGATTA CTAGATACAT ACACGTATAG  
2651 ATGGATGAAG TGGAGAGAGA TTAGAGATAA GTAGTTATAT GAATTTTGTG  
2701 AAACACACTT AAGACATATG TTCAAACATA CTGCTATTAT GTATGAAATA  
2751 TTGAGTTTTA ACGGTTTAAA ACACATATTC TTTTAATTAG AATGTAATAA  
2801 TGTGATATCT TGTGTGAAAA TTTAATTACA TCTAATATAA CGGTGTGATT  
2851 AGATTGTATG TTGGATAACA TGCCCATCGG TTGGCTTATT TAGGGAATAA  
2901 GCCAAATGGT ATATTTGCAA ACGAAAAATA ATTTGTAAAT AAAACTTTTA  
2951 TGTATGTATT CTTAACGATC TAGCAGCAAA GGCTGAAAAA TAAACTTCGA  
3001 TGAAAAATCT CAAAATCAAC TCTTAAAAAT TAAATTTTGG CTTATAAGTA  
3051 TAGTTCCTAA CTAGTTTAGA AGAAAAATA TTTAAAGCGG GGAAGAGGAA  
3101 AAGGAATAAA CTAATAGCTA AATTATTGCA TGCATGTAGC GATTTGAGGA  
3151 CGACCGAGTT GTTTGTCTG GATCAGCCGA CCGAGACAGA GCAATCTTCT  
3201 TTAATCATAA ATAACCAGAA AAACCATACC AGTTCATCAC AATGGACCGA  
3251 GTCAGAGTCA TTACATATTT TTCAATTGTTG CGCACAGGAT TCACCATGTT  
3301 CTTATGGGAA ATATTTTAA CTCTCAAATG GTTATGATTT TGAACCTCTCA  
3351 TTTTGTAGAG AGAATTAACA AGCGAGCGAG CAATCAGGCC AAAAAGGGAG  
3401 AAAGAAAAAT ATTTTGTFTA ATTTTTTTTT AAGGTAGGGT GGAGGAGTCA  
3451 TTACATGATT TTTTTTTATA TTCCCTCGTT GATTATATGC TGTTCAAATG  
3501 GTTATGATTT TTTTAAAAGA TAACAACAAT ACAAATTAGT ATGTGATAGA  
3551 TCATTTACAG AGCATATAGG ATTAAATTTA ACTTCTGTAA ATTACAAAAC  
3601 AAACAAGTTT AACTGTTAAT ATACATTAAA TTTGTTTTTT TCAACTTAGG  
3651 AATTGAATTT TATGTATATA TTTGTAAAAT GATATATTAA TTTATTTTTT  
3701 TAAAAAATA ATTATTTAGA TAACACGCAA ACTAGAAAAC CACCGCAGAA  
3751 GTTCTCATAT TTCTGTCTCT ATCTGCACTT GCAGAGCTGA TGCTGCTGGG  
3801 CTTCATATCC CTGCTTCTCA CCGTGGCACA GGCGCCCATC TCCAAGATCT

## FIGURE 8 cont'd

3851 GCATCCCCAA GTCGGCTGCC AACATCTTGT TGCCGTGCAA GGCAGGCCAA  
3901 GATGCCATCG AAGAAAGAAG CAGCAAGTGG TCGCCGGTCC TTGGCCGGCG  
3951 CCGGCGGCGG GGACTACTGC TCGAAATTCG ATGTGAGAAT AACACCAGCT  
4001 GCCGGCAAGC ACAACCTCGA TGCAATAACT AATTTAAC TAATTGATTT  
4051 TTCTTGGGTT TTCTGCAGGG CAAGGTGGCG CTGATGTCGG CAAAGAGCAT  
4101 GCACCAGCTG CACATTTTCA TCTTCGTGCT CGCCGTGTTT CATGTTACCT  
4151 ACTGCATCAT CACCATGGGT TTAGGGCGCC TCAAAGTGAG TTTGTCGTTC  
4201 TGTCCTTCAT GCACATGTTT TCTCTAGTTC TAGCAAGATT GTCAGTCCTT  
4251 CAAATGGATT GTTTCGACAA GAAACCCAAT TTATTAATTT GCCAGTAAAT  
4301 ATATAATAAT TGATCTTTCT TGGTTTTAGA TGAAGAAATG GAAGAAGTGG  
4351 GAGTCACAGA CCAACTCATT GGAGTATCAG TTCGCAATCG GTAGTGAATT  
4401 AAGAATCTCC CTAACATTTT CATTTTCAGAA CCTTTATGAT AATGTCTTGA  
4451 AAGAGGAGGA GCAAATCAGC TGAAAAATAT GATCGATCCA TGCAGATCCT  
4501 TCACGATTCA GGTTACGCA TCAGACGTCG TTCGTGAAGC GGCATCTGGG  
4551 ATCATTTCTA AGCACCCCTG GGCTCAGATG GATCGTGAGT TATCAATCTC  
4601 CGAATACATG CTTGTTTTTT ATTCTTGCAA CTGGCCTAGC TGTTCCAATT  
4651 CAATCCATAT TTTTGTAAAA AAAAAATATT CATGCCGTGT TTGTTGTTAG  
4701 GTAGCATTCT TCAGGCAGTT CTTTGGGTCC GTCACCAAGG TGGACTACCT  
4751 GACCATGCGG CAAGGCTTCA TCAATGTATA TACTAATCAA ACCTGACCAA  
4801 TTCAACATTG ATGATGCAA CAGAGACCAG GTTTTTTTTT TCGAGTGTGC  
4851 ATTGAGTAAT GGTTTTAGCT TCTCTCTTTT TGCAGGCGCA TTTGTCGCAG  
4901 AATAGCAAGT TCGACTTCCA CAAATACATC AAGAGGTCTT TGGAGGACGA  
4951 CTTCAAAGTT GTCGTTGGCA TCAGGTCCGT CCTCGCTTTA TTAATTATAG  
5001 GACTCTTATA TTCAACATTT TTTTATAAAA GAAACATATT TAGTCTCCAG  
5051 TTGTGTATGT GTATGTGGAT CTTGACACAT TTGGCTGGTT TTGCAGCCTC  
5101 CCTCTGTGGT TCGTCGGAAT CCTGTACTC TTCTCGATA TCCACGGTAA  
5151 TCCTTGTCCT ATTTCAATCT TTTTTTACT CTCAAACCT TGTCTGAAT  
5201 TGGTCTTATA ATCACCATCG ATTTTTTTT AACTTTTTTC CCGCGTGTAG  
5251 GTCTTGGCAC ACTTATTTGG ATCTCTTTTG TTCTCTCAT CGTAAGAGCG  
5301 AAATTTCCCT GTCCAAAGAA ACAGTTAACA TAATTAATTA TGCTTTAATT  
5351 TATCATGAAA ATTAATATGA TCATATAACT AATGAACAAA CATTATGTG  
5401 AATGCCACCG TTGTCTCAGA TCGTCTTGT AGTTGGGACC AAGCTAGAGA  
5451 TGGTGATCAT GGAGATGGCC CAAGAGATAC AGGACAGGGC CACTGTGATC  
5501 CAGGGAGCAC CTATGGTTGA ACCAAGCAAC AAGTACTTCT GGTCAACCG  
5551 CCCTGACTGG GTCTTGTTC TACATACACCT GACACTCTTC CATGTACATG  
5601 TTTAAAACCT AAACCTTGCT GCTCAACTAC AAATAGTACT TTATCTTTCA  
5651 CAATTAACAC CTAATTAAC AACATAGCAT CCATCCATTT GTGGCTACTG  
5701 ATCGATGGGA CGACGGATCG ATCATCACCA GAACGCATTT CAGATGGCGC  
5751 ATTCGTATG GACTATGGTG TGTATGCTAC TTGCTTAGTT GTTGCCATTA

## Figure 8 cont'd

5801 TCAGTTCTTA AGCAAATTAA GTGTGATGCA TGCCTGACT AATGAGACAA  
5851 AAAATGACAC AGCTTGTTCA TCGATCTGGT TGTTTTGTGT GTGACAGGCA  
5901 ACACCTGGTC TGAAGAAATG CTTCCATGAA AATATTTGGC TGAGCATCGT  
5951 GGAAGTCATT GTGGGGATCT CTCTTCAGGT GCTATGCAGC TACATCACCT  
6001 TCCCGCTCTA CGCGCTCGTC ACACAGGTGA ACAAGCCATT CACAAATTCT  
6051 ATTAGCCGTT TCTTAATTGA TGACACTGTT AATTTTATA CACACGTTTT  
6101 GACCATTTGT CTTATTAAAA ATATTTATGT AATTATCATT TGAGTTGTTT  
6151 TATCACTAAA AGTACTTTTT AAATAATTAA TATTTTGCAT TTGTACAATT  
6201 CTTTTAATAA GATAATGGTC AAACATGTGT CCAAAAGTTA ACAGCATCAT  
6251 CTATTAAGAA AAGGAGGGGT TTTTTTTTTT TGGAAATTTG CAAAATTTGT  
6301 TCAAAATCAG TCCAAAACCT TTTTTTTTTT CGAAATTTCA GTTTCCTAC  
6351 CAGTCCCAT AAAATGTCTT TTCTTTATTT CCACAAGATT GAACCCATGA  
6401 GATGCCCTTT GTGTTGGTAT GTGTTTGGC CATCACTTGC AGATGGGATC  
6451 GAACATGAAG AAGACAATTT TCGAGGAGCA AACGATGAAG GCGCTGATGA  
6501 ACTGGAGGAA GAAGGCGATG GAGAAGAAGA AGGTCCGGGA CGCCGACGCG  
6551 TTCCTGGCGC AGATGAGCGT CGACTTCGCG ACGCCGGCGT CGAGCCGGTC  
6601 CGCGTCGCGG GTGCACCTGC TGCAGGTCAC AGGGCGGGTC GGACGCCCCG  
6651 CGAGCCCAAT CACGETGGCC TCACCACCGG CACCGGAGGG GACATGTACC  
6701 CGGTGCCGGC GCGCGCTCGG TCTCGCCAGC TGCTAGACGA CCCGCCGGAC  
6751 AGGAGGTGGA TGGCATCCTC GTCGGCCGAC ATCGCCGATT CTGATTTTTC  
6801 CTTCAGCGCA CAACCGTGAC GGGGGCGATC GCTTTCTGTA TTGATGCTGT  
6851 ACCAAACATA GGAGTTTAAT ATATATATAA TTGTTACGGT AAAATCTAAT  
6901 TATTGTGCGC GCACTTATAT TAGTCTTATA GCGCGACTGG TTCGTGATTA  
6951 GACAAGGTGA TGCATGCTGT TTAGTTATAA AGGATATCAG CGCAGCTAAA  
7001 AAAACTTACT CCCTACTTAA TAGATGACCT CGTTGATTTT TAACATTATT  
7051 CGTCTTATTT AAAAATTGA TGCAAATGTT TAAAACATAA ATCATGCTTA  
7101 AAGTACTTTT AGTGATAAAA CACTTACAA CAAAATAAAT TATAGTTACC  
7151 TAATTTTTTT TAATAAATCG AATGG

0972237.1.12300

1 TTATACCATG TGAGAAAGGC TGGAAGCATA TGCTCTTAGC AGGGACGCGT  
51 GCATGTTTAT ATAGGAGGCA TAAGCCGAAG AGATATACAT GAGGAGAGGT  
101 TTAAGATCAG TCTATCTTAT TTACAGTTTA AACACAAGGA GATAGAAAGA  
151 GATCCTAACC TACACATGTT ATACAAGTCA CGTATAATAC AAGAGTTATT  
201 TCGTCTAACA CCTTCCCCTC TGATATGATA AGTCGCCGGG AGAGAGAGAG  
251 AGTGTGTGGC TGCCCTCGCT GCACTGCACG CACATGTTTA CTTCTCCGAC  
301 TGAAACCACG GTGAAACCGG CGGCGGTGTC GCACTCCCCT GACTTTCCTC  
351 GCCGGGGTCC CGTCCGGACA ATTAAACCGT CTGTACCTGC CGGGCGTCGA  
401 CCCGATCGTG ATGTGGCGCC GCTTTGTCTG CAGCGAGCTG CGTGGCCGAT  
451 GGCAACAAAA CTGCGGTCAC ATACATGCAT ACCCCGCATA CCCCGACGCT  
501 CACCAGTAAG TAGGCTGTGG TGCGGCACCA CGGGCTCGCC GCCATTTCATG  
551 CCATGCATGG GCCACCCGCC GCGGAAACCG CGGCGCTGCT GCCTGCCACC  
601 CCGCCGCCGT TGACGAAGAC TTCGCCCGGC CATCCATAAA AGCATGCATG  
651 GCTTGCTCTC ACCGGTCCGG CCACACACAC CACACTTCAC TTCGCCATTC  
701 GCACCACCGA GAGCGTAGCG TAACGTGTGT TTGAAGTCCT ACCATTAATT  
751 TTGCTGGATC GATGGCTGGG CCGGCGGGAG GTCGGGAGCT GTCGGACACG  
801 CCGACGTGGG CGGTGGCGGT AGTCTGCGCC GTCATGATAC TCGTCTCCGT  
851 CGCCATGGAG CACGCGCTCC ACAAGCTCGG CCACGTACGT GCTCTCGGTT  
901 CACTAGTGCT TAACTGTTTT TGATGTTTTT GGGCGTGTTC GGTAGCCTGC  
951 ATGGAGAGTG TATGAGCCCA AAAGTTCCCT CCCCGACCCA CTTTTCGCTG  
1001 TTTGGTAGGG TGTATGGGCT GAGGAGAGCA TGCATCAACT GATGCAAAAA  
1051 GGGCCTCAGC ATAGCTGAGC CCAGCACCCC CGCAGAGGCG AGCTGAGGCG  
1101 AGTTATGCTG AGCCCATGCA CCTTCGCCCC GTCGCCCCGT CGCCCCGTGC  
1151 CTCCCCCCT GCACCTCTTC CTCTCCCTC TTCCTACCAA ACACAGTCTC  
1201 ATCCAAACAT GTAACAACAC ATGCATGACC ACCAAACAAC TGAAGATGAA  
1251 TGTATTTCATC ATGTCTATAC TTACCATGCA TCAACAGGGA ACAACTATGC  
1301 TAGGGTGAGA ACAGCTGCCA AACACACCCG TGCACCTACT CATGCTGTGC  
1351 CGGCGCTGGC GTACGTGTGC AGTGGTTCCA CAAGTGGCGC AAGAAGGCCC  
1401 TGGGGGAGGC GCTGGAGAAG ATGAAGGCGG AGCTCATGCT GGTGGGCTTC  
1451 ATATCCCTGC TCCTCATCGT CACGCAGGAT CCCGTCTCCA GGATCTGCAT  
1501 CTCCAAGGAG GCCGGCGAGA AGATGCTCCC GTGCAAGCCT TACGACGGCG  
1551 CCGGCGGTGG CAAAGGCAAG GACAATCACC GGAGGCTTCT CTGGCTCCAA  
1601 GGCGAGAGCG AGACCCACCG CCGGTTCCCTG GCTGCCCCGG CCGGAGTGGA  
1651 CGTCTGCGCC AAACAGGTGA GCACCTAGCG TCGCCACAAA CCACAACTA  
1701 GCTAATGAGC ATGGACCTGA ATTTCTTCTC TTCTTGGCTT GGCTTGACTA  
1751 AATTGGTTGT GCAGGGCAAG GTGGCGCTGA TGTCAGCGGG AAGCATGCAC  
1801 CAACTGCACA TATTCATCTT CGTGCTCGCC GTCTTCCACG TCTTGTACAG  
1851 CGTCGTCAAC ATGACCCTAA GCCGTCTCAA AGTGAGCATC ATACTCGAGC

00927 22260

1901 TGTTTGTCAA TAATCCTTGG TTTCCAATCC AATTCCAAAG CTGGCACTGA  
1951 TCCTGCTCCG GCTTCCTGCA GATGAAGCAA TGGAAGAAGT GGGAGTCCGA  
2001 GACCGCCTCG CTGGAGTATC AGTTCGCGAA TGGTCAGCTT CAACTTTTCT  
2051 TACTGAAACC GGATGCATTT ACAACAAACG CACGCACGAT CAATCATCAC  
2101 AGTGTGAGCC GATACGTTGA ACCGATTGAA TCCTCGCAGA TCCATCGCGG  
2151 TGCCGGTTCA CGCACCAGAC GACGTTGGTG AGGCGGCACC TGGGCCTCTC  
2201 CAGCACCCCC GCGCTCAGAT GGGTGGTGGC CTTCTTCAGG CAGTTCTTCA  
2251 CGTCGGTGAC CAAGGTGGAC TACCTGACCT TGCGGCAGGG CTTTCATCAAC  
2301 GCGCATCTCT CGCAGGGCAA CAGGTTGCGAC TTCCACAAGT ACATCAAGAG  
2351 GTCGTTGGAG GACGACTTCA AAGTCGTCGT CCGCATCAGG TACGCGCCAT  
2401 TCCTTTCTCT GCACAAATTA ATACATCCAC CACCACATAG GTAGATAGAT  
2451 AGATCGATAG ATAGATTATA CAAGTGCCGG TACGTACGTA CGTCTCATAT  
2501 GATCTTGACA CATCTGTCCT CTTGCCGCGAG TCTCAAGCTC TGGTTCGTGG  
2551 CGGTCTCAT CCTCTTCCTT GATTTTCGAG GTAGCCGCCT TGTCCATGCC  
2601 CTGCTCGCCC TCTCCTCCGC TTCTCTCCAT AATTTGTGAA CTTGTCCCGT  
2651 ATATAACCAC ACCACCGTCG TCTTCTCGCA GGGATCGGCA CTCCTCTCTG  
2701 GATGTCCGTG GTTCTCTCTG TGGTAAGTCC ACAATTTGAA TAGACAACCT  
2751 GTCCAATTGT GATGTACAGT ACCTCCAAAC TTAATTAACA TGTCATTTGC  
2801 TGATGTCTTG CGTGTAACAT TAGATCCTCT TGTGGGTTGG GACCAAGCTG  
2851 GAGATGGTGA TCATGGAGAT GGCCCAGGAG ATCCATGACC GGGAGAGCGT  
2901 CGTCAAGGGT GCTCCCGCCG TCGAGCCCAG CAACAAGTAC TTCTGGTTCA  
2951 ACCGGCCTGA CTGGGTCCTC TTCCTCATGC ACCTCACACT CTTCCAGAAC  
3001 GCGTTTCAGA TGGCTCATTT CGTGTGGACA GTGGTACGTA CAAGTACTTG  
3051 TCACTTCACT TAGGCTAACT CCAACAAACG ACCCCAAATT AATGGTCCGT  
3101 CGCGTCTGTT TGGGGTATGT TTGGGGTAAA CGGACACAAA ACTCAATCCA  
3151 ACGCGCGGTA GCAAACGAAC GTTTTTCGGT ACGTTTTTCGT CCGCTTTTCG  
3201 CCCATCCCAG CCCAAATTCG TTGACGTTGT TGCATCGCAG GCCACGCCCC  
3251 GCTTGAAGAA ATGCTACCAC GAGAAAATGG CAATGAGCAT CGCCAAGGTC  
3301 GTGCTGGGGG TAGCCGCCCA GATCTTGTGC AGCTACATCA CCTTCCCGCT  
3351 CTACGCGCTC GTCACGCAGA TGGGCTCACA CATGAAGAGA AGCATCTTCG  
3401 ACGAGCAGAC GGCCAAGGCG CTGACCAACT GGCGAAAGAT GGCCAAGGAG  
3451 AAGAAGAAGG CCCGAGACGC GGCCATGCTG ATGGCGCAGA TGGGCGGCGG  
3501 CGCGACGCCG AGCGTCGGCT CGTCGCCGGT GCACCTGCTC CACAAGGCCG  
3551 GGGCGCGGTC CGACGACCCC CAGAGCGTGC CGGCGTCCCC GAGGGCCGAG  
3601 AAGGAAGGCG GCGGCGTGCA GCATCCGGCG CGCAAGGTAC CTCCTTGTGA  
3651 CGGGTGGAGG TCGGCCTCGT CGCCGGCGCT CGACGCTCAC ATCCCCGGTG  
3701 CAGATTTTGG CTTTCAGCACG CAACGTTGAC CGATCAGACA AGTTCTTTT  
3751 TTTTTCGGTG AATAGAAGCG TATCATTTCA TTGATAGACA GTAGAAATTA  
3801 CAGGAATGGC TGTCTACTA CTATGTACAC AAGGGCACAG CAAAGGATCA



## Figure 10

1 ATGGCAGGTG GGAGATCGGG ATCGCGGGAG TTGCCGGAGA CGCCGACGTG  
 51 GGC~~GGTGGCC~~ GTCGTCTGCG CCGTCCTCGT GCTCGTCTCC GCCGCCATGG  
 101 AGCACGGCCT CCACAACCTC AGCCATAAAA CCACCGCAGA AGTTCTCATA  
 151 TTTCTTGTC TATCTGCACT TGCAGAGCTG ATGCTGCTGG GCTTCATATC  
 201 CCTGCTTCTC ACCGTGGCAC AGGCGCCCAT CTCCAAGATC TGCATCCCCA  
 251 AGTCGGCTGC CAACATCTTG TTGCCGTGCA AGGCAGGCCA AGATGCCATC  
 301 GAAGAAGAAG CAGCAAGTGG TCGCCGGTCC TTGGCCGGCG CCGCGGGCGG  
 351 GGACTACTGC TCGAAATTCTG ATGGCAAGGT GGCCTGATG TCGGCAAAGA  
 401 GCATGCACCA GCTGCACATT TTCATCTTCG TGCTCGCCCT GTTCCATGTT  
 451 ACCTACTGCA TCATCACCAT GGGTTTAGGG CGCCTCAAAA TGAAGAAATG  
 501 GAAGAAGTGG GAGTCACAGA CCAACTCATT GGAGTATCAG TTCGCAATCG  
 551 ATCCTTCACG ATTCAGGTTT ACGCATCAGA CGTCGTTCTG GAAGCGGCAT  
 601 CTGGGATCAT TCTCAAGCAC CCCTGGGCTC AGATCGGATCG TAGCATTTCTT  
 651 CAGGCAGTTC TTTGGGTCCG TCACCAAGGT GGACTACCTG ACCATGCGGC  
 701 AAGGCTTCAT CAATGCGCAT TTGTGCGAGA ATAGCAAGTT CGACTTCCAC  
 751 AAATACATCA AGAGGTCTTT GGAGGACGAC TTCAAAGTTG TCGTTGGCAT  
 801 CAGCCTCCCT CTGTGGTTCG TCGGAATCCT TGTACTCTTC CTCGATATCC  
 851 ACGGTCTTGG CACACTTATT TGGATCTCTT TTGTTCTCT CATCATCGTC  
 901 TTGTTAGTTG GGACCAAGCT AGAGATGGTG ATCATGGAGA TGGCCCAAGA  
 951 GATACAGGAC AGGGCCACTG TGATCCAGGG AGCACCTATG GTTGAACCAA  
 1001 GCAACAAGTA CTTCTGGTTC AACCGCCCTG ACTGGGTCTT GTTTTTCATA  
 1051 CACCTGACAC TCTTCCATAA CGCATTTTCAG ATGGCGCATT TCGTATGGAC  
 1101 TATGGCAACA CCTGGTCTGA AGAAATGCTT CCATGAAAT ATTTGGCTGA  
 1151 GCA~~TCGTGGA~~ AGTCATTGTG GGGATCTCTC TTCAGGTGCT ATGCAGCTAC  
 1201 ATCACCTTCC CGCTCTACGC GCTCGTCACA CAGATGGGAT CGAACATGAA  
 1251 GAAGACAATT TTCGAGGAGC AAACGATGAA GGCCTGATG AACTGGAGGA  
 1301 AGAAGGCGAT GGAGAAGAAG AAGGTCCGGG ACGCCGACGC GTTCTTGCGG  
 1351 CAGATGAGCG TCGACTTCGC GACGCCGGCG TCGAGCCGGT CCGCGTCGCC  
 1401 GGTGCACCTG CTGCAGGTCA CAGGGCGGGT CGGACGCCCG CCGAGCCCCA  
 1451 TCACGGTGGC CTCACCACCG GCACCGGAGG AGGACATGTA CCCGGTGCCG  
 1501 GCGGCGGCTG CGTCTCGCCA GCTGCTAGAC GACCCGCCGG ACAGGAGGTG  
 1551 GATGGCATCC TCGTCGGCCG ACATCGCCGA TTCTGATTTT TCCTTCAGCG  
 1601 CACAACGGTG A

093274422460

22/28

## Figure 11

1 ATGGCTGGGC CGGCGGGAGG TCGGGAGCTG TCGGACACGC CGACGTGGGC  
51 GGTGGCGGTA GTCTGCGCCG TCATGATACT CGTCTCCGTC GCCATGGAGC  
101 ACGCGCTCCA CAAGCTCGGC CACTGGTTCC ACAAGTGGCG CAGAGAGGCC  
151 CTGGGGGAGG CGCTGGAGAA GATGAAGGCG GAGCTCATGC TGGTGGGCTT  
201 CATATCCCTG CTCCTCATCG TCACGCAGGA TCCCGTCTCC AGGATCTGCA  
251 TCTCCAAGGA GGCCGGCGAG AAGATGCTCC CGTGCAAGCC TTACGACGGC  
301 GCCGGCGGTG GCAAAGGCAA GGACAATCAC CGGAGGCTTC TCTGGCTCCA  
351 AGGCGAGAGC GAGACCCACC GCGGTTCCTT GGCTGCCCCG GCCGGAGTGG  
401 ACGTCTGCCG CAAACAGGGC AAGGTGGCGC TGATGTCAGC GCGAAGCATG  
451 CACCAACTGC ACATATTTCAT CTTCGTCTTC GCCGTCTTCC ACGTCTTGTA  
501 CAGCGTCGTC ACCATGACCC TAAGCCCTCT CAAATGAAG CAATGGAAGA  
551 AGTGGGAGTC GGAGACCGCC TCGCTGGAGT ATCAGTTCGC GAATGATCCA  
601 TCGCGGTGCC GCTTCACGCA CCAGACGACG TTGGTGAGGC GGCACCTGGG  
651 CCTCTCCAGC ACCCCCGGCG TCAGATGGGT GGTGGCCTTC TTCAGGCACT  
701 TCTTCACGTC GGTGACCAAG GTGGAATACC TGACCTTGCG GCAGGGCTTC  
751 ATCAACGCGC ATCTCTCGCA GGGCAACAGG TTCGACTTCC ACAAGTACAT  
801 CAAGAGGTCG TTGGAGGACG ACTTCAAAGT CGTCGTCCGC ATCAGTCTCA  
851 AGCTCTGGTT CGTGGCGGTC CTCATCCTCT TCCTTGATTG CGACGGGATC  
901 GGCACCTCTC TCTGGATGTC CGTGGTTCCT CTCGTGATCC TCTTGTGGGT  
951 TGGGACCAAG CTGGAGATGG TGATCATGGA GATGGCCCAG GAGATCCATG  
1001 ACCGGGAGAG CGTCGTCAAG GGTGCTCCCG CCGTCGAGCC CAGCAACAAG  
1051 TACTTCTGGT TCAACCGGCC TGAATGGGTC CTCTTCCTCA TGCACCTCAC  
1101 ACTCTTCCAG AACGCGTTTC AGATGGCTCA TTTCGTGTCG ACAGTGGCCA  
1151 CGCCCGGCTT GAAGAAATGC TACCACGAGA AAATGGCAAT GAGCATCGCC  
1201 AAGGTGCTGC TGGGGGTAGC CGCCAGATC TTGTGCAGCT ACATCACCTT  
1251 CCCGCTCTAC GCGCTCGTCA CGCAGATGGG CTCACACATC AAGAGAAGCA  
1301 TCTTCGACGA GCAGACGGCC AAGGCGCTGA CCAACTGGCG AAAGATGGCC  
1351 AAGGAGAAGA AGAAGGCCCG AGACGCGGCC ATGCTGATGG CGCAGATGGG  
1401 CGGCGGCGCG ACGCCGAGCG TCGGCTCGTC GCCGGTGCAC CTGCTCCACA  
1451 AGGC CGGGG GCGGTCCGAC GACCCCAAGA GCGTCCCGGC GTCCCGGAGG  
1501 GCGGAGAAGG AAGGCGGCGG CGTGCAGCT CCGGCGCGCA AGGTACCTCC  
1551 TTGTGACGGG TGGAGGTCGG CCTCGTCGCC GCGGCTCGAC GCTCACATCC  
1601 CCGCTGCAGA TTTTGGCTTC AGCACGCAAC GTTGA

097637.12800



0972377.1300

1 GTTGGTACAT AAAAGACTCT TCCTTTGTCT GTTTTTGT TT CCCAGATTCA  
51 TCTTTACTTA TTGACTAAAT TCTCTCTGGT GTGAGAAGTA AAATGGGTCA  
101 CGGAGGAGAA GGGATGTCGC TTGAATTCAC TCCGACGTGG GTCGTCGCCG  
151 GAGTTTGTAC GGTCATCGTC GCGATTTTAC TGGCGGTGGA GCGTTTGCTT  
201 CACTATTTTCG GTACTGTTCT TAAGAAGAAG AAGCAAAAAC CCCTTTACGA  
251 AGCCCTTCAA AAGGTTAAAG AAGAGCTGAT GTTGTTAGGG TTTATATCGC  
301 TGTTACTGAC GGTATTCCAA GGGCTCATTT CCAAATTCTG TGTGAAAGAA  
351 AATGTGCTTA TGCATATGCT TCCATGTTCT CTCGATTCAA GACGAGAAGC  
401 TGGGGCAAGT GAACATAAAA ACGTTACAGC AAAAGAACAT TTTCAGACTT  
451 TTTTACCTAT TGTGGAACC ACTAGGCGTC TACTTGCTGA ACATGCTGCT  
501 GTGCAAGTTG GTTACTGTAG CGAAAAGGGT AAAGTACCAT TGCTTTTCGT  
551 TGAGGCATTG CACCATCTAC ATATTTTCAT CTTCGTCCTC GCCATATCCC  
601 ATGTGACATT CTGTGTCCTT ACCGTGATTT TTGGAAGCAC AAGGATTCAC  
651 CAATGGAAGA AATGGGAGGA TTCGATCGCA GATGAGAAGT TTGACCCCGA  
701 AACAGCTCTC AGGAAAAGAA GGGTCACTCA TGTACACAAC CATGCTTTTA  
751 TTAAAGAGCA TTTTCTTGGT ATTGGCAAAG ATTCAGTCAT CCTCGGATGG  
801 ACGCAATCCT TTCTCAAGCA ATTCTATGAT TCTGTGACGA AATCAGATTA  
851 CGTGACTTTA CGTCTTGGTT TCATTATGAC ACATTGTAAG GGAAACCCCA  
901 AGCTTAATTT CCACAAGTAT ATGATGCGCG CTCTAGAGGA TGATTTCAAA  
951 CAAGTTGTTG GTATTAGTTG GTATCTTTGG ATCTTTGTCTG TCATCTTTTT  
1001 GCTGCTAAAT GTTAACGGAT GGCACACATA TTTCTGGATA GCATTTATTC  
1051 CCTTTGCTTT GCTTCTTGCT GTGGGAACAA AGTTGGAGCA TGTGATTGCA  
1101 CAGTTAGCTC ATGAAGTTGC AGAGAAACAT GTAGCCATTG AAGGAGACTT  
1151 AGTGGTGAAA CCCTCAGATG AGCATTTCTG GTTCAGCAA CCTCAAATTG  
1201 TTCTCTACTT GATCCATTTT ATCCTCTTCC AGAATGCTTT TGAGATTGCG  
1251 TTTTCTTTTT GGATTTGGGT TACATACGGC TTCGACTCGT GCATTATGGG  
1301 ACAGGTGAGA TACATTGTTC CAAGATTGGT TATCGGGGTC TTCATTCAAG  
1351 TGCTTTGCAG TTACAGTACA CTGCCTCTTT ACGCCATCGT CTCACAGATG  
1401 GGAAGTAGCT TCAAGAAAGC TATATTCGAG GAGAATGTGC AGGTGGTCTT  
1451 TGTTNGTTGG GCACAGAAAG TGAAACAAA GAGAGACCTA AAAGCTGCAG  
1501 CTAGTAATGG AGACGAAGGA AGCTCTCAGG CTGGTCCTGG TCCTGATTCT  
1551 GGTTCTGGTT CTGCTCCTGC TGCTGGTCCT GGTGCAGGTT TTGCAGGAAT  
1601 TCAGCTCAGC AGAGTAACAA GAAACAACGC AGGGGACACA AACAATGAGA  
1651 TTACACCTGA TCATAACAAC TGAGCAGAGA TATTATCTTT TCCATTTAGA  
1701 GGATCATCAT CAGATTTTAG CTTCAAGGTC CGGTTTGTG GTTTATACAT  
1751 AAGTTATAGT GACTTGATTT TTTTGTTTTG TTACAAAGTT ACCATCTTTG  
1801 GATTAGAATT GGGAAATTGA ATCTGTTTGT ATATTGTATT ATTTGGAACA  
1851 TTGTGGATGC CCATGGATAT GTTCTGTTC

1 MAGGRSGSRE LPETPTWAVA VVCAVLVLVS AAMEHGLHNL SHKTTAEVLI  
51 FLVLSALAE LMLGFISLLL TVAQAPISKI CIPKSAANIL LPCKAGQDAI  
101 EEEAASGRRS LAGAGGGDYC SKFDGKVALM SAKSMHQLHI FIFVLAVFHV  
151 TYCIITMGLG RLKMKKKWKW ESQTN SLEYQ FAIDPSRFRF THQTSFVKRH  
201 LGSFSSTPGL RWIVAFFRQF FGSVTKVDYL TMRQGFINAH LSQNSKFDFH  
251 KYIKRSLEDD FKVVVGISLP LWFVGILVLF LDIHGLGTLI WISFVPLIIV  
301 LLVGTKLEMV IMEMAQEIQD RATVIQAPM VEPSNKYFWF NRPDWVLFPI  
351 HLTLFHNAFQ MAHFVWTMAT PGLKKCFHEN IWLSIVEVIV GISLQVLC SY  
401 ITFPLYALVT QMGSNMKTII FEEQTMKALM NWRKKAMEKK KVRDADAFLA  
451 QMSVDFATPA SSRASPVHL LQVTGRVGRP PSPITVASPP APEEDMYFVP  
501 AAAASRQLLD DPPDRRWMA'S SSADIADSD F SFAQR\*

1	MAGPAGGREL	SDTPTWAVAV	VCAVMILVSV	AMEHALHKLK	HWFHKWRKKA
51	LGEALEKMKA	ELMLVGFISL	LLIVTQDPVS	RICISKEAGE	KMLPCKPYDG
101	AGGGKGKDNH	RRLWLQGES	ETHRRFLAAP	AGVDVCAKQG	KVALMSAGSM
151	HQLHIFIFVL	AVFHVLYSVV	TMTLSRLKMK	QWKKWESETA	SLEYQFANDP
201	SRCRFTHQTT	LVRRLGLSS	TPGVRVWVAF	FRQFTSVTK	VDYLTLRQGF
251	INAHLSQGNR	FDFHKYIKRS	LEDDFKVVVR	ISLKLWFVAV	LILFLDFDGI
301	GTLLWMSVVP	LVILLWVGTK	LEMVIMEMAQ	EIHDRESVVK	GAPAVEPSNK
351	YFWFNRPDWV	LFLMHLTLFQ	NAFQMAHFVW	TVATPGLKKC	YHEKMAMSIA
401	KVVLGVAAQI	LCSYITFPLY	ALVTQMGSHM	KRSIFDEQTA	KALTNWRKMA
451	KEKKKARDAA	MLMAQMGGGA	TPSVGSSPVH	LLHKAGARSD	DPQSVPASPR
501	AEKEGGGVQH	PARKVPPCDG	WRSASSPALD	AHIPGADFGF	STQR*

**Figure 15**

1 MCHGGEGMSL EFTPTWVVAG VCTVIVAISL AVERLLHYFG TVLKKKKOKP  
51 LYEALQKVKE ELMLLGFISL LLTVFQGLIS KFCVKENVLM HMLPCSLDSR  
101 REAGASEHKN VTAKCHFQTF LPIVGTTRRL LAEHAAVQVG YCSEKGVPL  
151 LSLEALHHLH IFIFVLAISH VTFCVLTVIF GSTRIHQWKK WEDSIADKEF  
201 DPETALRKRR VTHVHNHAFI KEHFLGIGKD SVILGWTQSF LKQFYDSVTK  
251 SDYVTLRLGF IMTHCKGNPK LNFHKYMMRA LEDDFKQVVG ISWYLWIFVV  
301 IFLLLNVNGW HTYFWIAFIP FALLLAVGTK LEHVIAQLAH EVAEKHVAIF  
351 GDLVVKPSDE HFWFSKPQIV LYLIHFILFQ NAFELAFFFW IWVTYGFDESC  
401 IMGQVRYIVP RLVIGVFIQV LCSYSTLPLY AIVSQMGSSF KKAILEENVQ  
451 VGLVGWAQKV KQKRDLCAAA SNGDEGSSQA GPGPDSGSGS APAAGPGAGF  
501 AGIQLSRVTR NNAGDTNNEI TPDHNN\*

09723377 112500

00327 4466460

FIGURE 16

Hvml0-H1	MAGPAG..GR	ELSDTPTTAV	AVVCAVMVLV	SVAMEH ALHK	LGHWFHKKRK	KALG EAL EKM
Mlo	MSDKKGV PAR	ELPETPTTAV	AVVFAAMVLV	SVLMEHGLHK	LGHWFQHRHK	KALWEAL EKM
Osmlo-H1m	MAGGRS..GSR	ELPETPTTAV	AVVCAVMVLV	SAAMEHGLHN	EGTFLKKTAKQ	EVLIFLWLSA
Atmlo-H1	..MGHGGEGM	SLFTPTTAV	AGVCTVIVAV	SLAMEHLLHY	EGTFLKKTAKQ	KPLYEALQKV
Consensus	M---G---R	EL--TPTTAV	AVVCAV-VLV	S-AMEH-LH-	LGH--K--	K-L-EAL-K-
Hvml0-H1	KAELMLVGFI	SLLLIVTQDP	MSKICISKE	AGEKMLPC	..KPYDGGAG	GKGKDNHRRRL
Mlo	KAELMLVGFI	SLLLIVTQDP	..IAKICISKE	AAADVMP	..K..	RGTEGRKPS
Osmlo-H1m	LAELMLVGFI	SLLLIVTQDP	..ISKICISKE	AAADVMP	..K..	EEAASGRRS
Atmlo-H1	KEELMLVGFI	SLLLIVTQDP	..ISKICISKE	AAADVMP	..K..	HKNVTAKEHF
Consensus	KAELML-GFI	SLLL-V-Q-P	..ISKICISKE	A--MLPC-	..K--A-	---
Hvml0-H1	LWLQGESETH	RRFLAAPAGV	DV.CAK.QGK	VALMSAGSMH	OLHIFIFVLA	VFHVLVSVVT
Mlo	KYVD.....	..LAGAGGG	..YCP.EGK	VALMSTGSLH	OLHIFIFVLA	VFHVTVSVIT
Osmlo-H1m	..	..LAGAGGG	..YCP.EGK	VALMSAGSMH	OLHIFIFVLA	VFHVTVSVIT
Atmlo-H1	QTFLPVGT	RRLLAEHAAV	QVGYCSEK	VALMSAGSMH	OLHIFIFVLA	VFHVTVSVIT
Consensus	---	..LA-	..GK	VALMS--S-H	OLHIFIFVLA	VFHVTV-V-T
Hvml0-H1	MTLSRLKMKQ	WKKWSESETH	LEYQFANDPS	RCRFT..HOT	TEVRRHLG.L	SSTPGV.RMV
Mlo	IALSRLKMKRT	WKKWSESETH	LEYQFANDPS	RCRFT..HOT	TEVRRHLG.L	SSTPGI.RMV
Osmlo-H1m	MGELRLKMKRT	WKKWSESETH	LEYQFANDPS	RCRFT..HOT	TEVRRHLG.L	SSTPGL.RMV
Atmlo-H1	VIEGSTTHQ	WKKWSESETH	LEYQFANDPS	RCRFT..HOT	TEVRRHLG.L	SSTPGL.RMV
Consensus	--L-RLKM--	WKKWE--T-S	LEYQFA-DP-	R-RFT--HOT	TEVRRHLG.L	SSTPGV.RMV
Hvml0-H1	VAFRQFFTS	VTKVVDYLTLR	QGFINAHLSQ	GNKFDPHKYI	KRSLEDDDFKV	VVRISLKLWF
Mlo	VAFRQFFTS	VTKVVDYLTLR	QGFINAHLSQ	GNKFDPHKYI	KRSLEDDDFKV	VVRISLKLWF
Osmlo-H1m	VAFRQFFTS	VTKVVDYLTLR	QGFINAHLSQ	GNKFDPHKYI	KRSLEDDDFKV	VVRISLKLWF
Atmlo-H1	QSEKQFFDS	VTKVVDYLTLR	QGFINAHLSQ	GNKFDPHKYI	KRSLEDDDFKV	VVRISLKLWF
Consensus	VAFRQFF-S	VTKVVDYLTLR	QGFINAHLSQ	N-KPDPHXYI	KRSLEDDDFKV	VVRISLKLWF
Hvml0-H1	VAVLILFLDF	DGIGTLLMMS	VVPLVILLLV	GTKLEMVIME	MAQELHDDRAS	VVKGAPAVEP
Mlo	VAVLILFLDF	DGIGTLLMMS	VVPLVILLLV	GTKLEMVIME	MAQELHDDRAS	VVKGAPAVEP
Osmlo-H1m	VAVLILFLDF	DGIGTLLMMS	VVPLVILLLV	GTKLEMVIME	MAQELHDDRAS	VVKGAPAVEP
Atmlo-H1	VAVLILFLDF	DGIGTLLMMS	VVPLVILLLV	GTKLEMVIME	MAQELHDDRAS	VVKGAPAVEP
Consensus	V--L-LFLD-	DGIGTLLMMS	VVPLVILLLV	GTKLEMVIME	MAQELHDDRAS	VVKGAPAVEP
Hvml0-H1	SNKFFWFNRPP	DMVLFLLHILT	LFQNAFQMAH	FVMTVATPGL	KKCYHHEKMMAM	SIKVVVVGVA
Mlo	SNKFFWFNRPP	DMVLFLLHILT	LFQNAFQMAH	FVMTVATPGL	KKCYHHEKMMAM	SIKVVVVGVA
Osmlo-H1m	SNKFFWFNRPP	DMVLFLLHILT	LFQNAFQMAH	FVMTVATPGL	KKCYHHEKMMAM	SIKVVVVGVA
Atmlo-H1	SNKFFWFNRPP	DMVLFLLHILT	LFQNAFQMAH	FVMTVATPGL	KKCYHHEKMMAM	SIKVVVVGVA
Consensus	SNK-FWF-RP	DMVLF-LHILT	LFQNAFQMAH	FVMTVATPGL	KKCYHHEKMMAM	SIKVVVVGVA

FIGURE 16 (CONT/D)

28/28

Hvml-o-H1	AQ	I L C S Y I T F F	P L Y A L V T Q M G	S H M K R S I F D E E	Q T A K A L T N W R	K M A K E K K K A R	D A A M L M A Q M G I
Mlo	LQ	F L C S Y M T F F	P L Y A L V T Q M G	S N M K R S I F D E E	A L T N W R	N T A K E K K K V R	D T D M L M A Q M I
Osmlo-Hlm	LQ	V L C S Y I T F F	P L Y A L V T Q M G	S N M K K T I F D E E	K A L M N W R	K K A M E K K K V R	D T D M L M A Q M S
Atmlo-H1	IQ	V L C S Y S T I F F	P L Y A L V S Q M G	S S F K K A I F E E E	N V Q V G L V G M A	Q K V E Q K R D L K	D A D A F L A Q M S
Consensus	-Q-	L C S Y - T F	P L Y A L V T Q M G	S - M K - - I F - E	Q T - K A L - N W R	- - A K E K K K - R	D A - - - A Q M -
Hvml-o-H1	GGA	T . . . . .	. P S V G S S P V	H L L H K K A G A R S	D D P Q S V P A S P	R A E K E G . . . .	. G G V Q H P A R K
Mlo	GDA	T P S R G S S	P M P S R G S S P V	H L L H K K G M G R S	D D P Q S A P T S P	R T Q Q E A R D M Y	. V V V A H P V H R
Osmlo-Hlm	VD	. . . . . F A T	P A S S R S A S P V	H L L Q . V T G R V	D D P P S P I T V A	S P P A P E E D M Y	P V P A A A S R Q
Atmlo-H1	SQA	G P G P D S G	S G S A P A A C P .	. . . . . G A G F A	G I Q L S R V T R N	N A G D T N N E I T	P P D H N N * . . .
Consensus	-A-	- - - - -	- - - - - S P V	H L L - - - - G R -	- - - - - P - S -	- - - - -	P - - - -
Hvml-o-H1	V . .	PCCDGGWR	SASSPALDAH	IPGADFGFS T	QR *		
Mlo	L . .	NNDRRRR	SASSSALEAD	IPSADEFSS Q	GR *		
Osmlo-Hlm	LLDD	PPDRRW	MASSS . . . A	IADSDFSPSA	QR *		
Atmlo-H1	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .		
Consensus	- - - - -	- P - D - - -	- ASS - - - - A -	I - - - DE - FS -	. . . . .		